HANDBOOK OF APPLIED PSYCHOLOGY

Volume One

VOLUME ONE

Handbook of

EDITED BY DOUGLAS H. FRYER, Ph.D. *New York University and Richardson, Bellows, Henry and Company, Inc.*

Contributors:

THEODORA M. ABEL · PAUL S. ACHILLES · LUTON ACKERSON G. LESTER ANDERSON · JOHN E. ANDERSON · ROSE G. ANDERSON DONALD E. BAIER · W. LESLIE BARNETTE, JR. · ROBERT J. BEITEL, JR. · ROGER M. BELLOWS · GEORGE K. BENNETT · ARTHUR L. BENTON · RALPH F. BERDIE · MARION A. BILLS · PHYLLIS BLANCHARD · FRANCIS F. BRADSHAW · STEUART HENDERSON BRITT · ALICE I. BRYAN · EMILY BURR · HAROLD E. BURTT JAMES W. CAMPBELL · MATTHEW N. CHAPPELL · PHILIP G. CORBY W. FRED COTTRELL · EDWINA A. COWAN · WARREN W. COXE D. H. DABELSTEIN · HERMAN DE FREMERY · ERNEST DICHTER JACK W. DUNLAP · BEATRICE J. DVORAK · JOSEPH W. EATON HORACE B. ENGLISH · GRANT FAIRBANKS · DEAN FARNSWORTH RAY FAULKNER · MARY FITE · JOHN C. FLANAGAN · THEODORE W. FORBES · ALBERT D. FREIBERG · DOUGLAS H. FRYER · LILLIAN M. GILBRETH · WILLIAM D. GLENN, JR. · WALLACE GOBETZ CARTER V. GOOD · WILLIAM S. GRAY · COLEMAN R. GRIFFITH THOMAS W. HARRELL · GEORGE W. HARTMANN · GORDON HENDRICKSON · EDWIN R. HENRY · L. P. HERRINGTON · REXFORD HERSEY · GERTRUDE HILDRETH · ROBERT HOPPOCK · WILLIAM A. HUNT · JAMES D. JACKSON · JOSEPH JASTROW · JOHN G.

NEW YORK · RINEHART

Applied Psychology

AND EDWIN R. HENRY, Ph.D. *Richardson, Bellows, Henry and Company, Inc., and Consultant, Standard Oil Company (New Jersey)*

JENKINS · RAYMOND A. KATZELL · ISABELLE KENDIG · FRANCIS
KEPPEL · HARRY D. KITSON · OTTO KLINEBERG · HERBERT E.
KRUGMAN · GEORGE LAWTON · KATE LEWIS · HENRY C. LINK
DARRELL BLAINE LUCAS · BERTHA M. LUCKEY · LILLIEN J. MARTIN
LOUIS WILLIAM MAX · T. R. McCONNELL · FERN McGRATH
HYMAN MELTZER · MILTON METFESSEL · BRUCE V. MOORE
EUGENE MYERS · LESTER N. MYER · GEORGE NAGY · T. ERNEST
NEWLAND · L. J. O'ROURKE · DONALD G. PATERSON · S. L. PRESSEY
GRACE RUBIN-RABSON · H. H. REMMERS · MARION W. RICHARDSON
EDWARD A. RUNDQUIST · GLADYS C. SCHWESINGER · CARL E.
SEASHORE · MORTON A. SIEDENFELD · LOWELL S. SELLING
DAVID SHAKOW · CARROLL L. SHARTLE · SADIE MYERS SHELLOW
E. DONALD SISSON · MARIE SKODAK · KINSLEY R. SMITH
DANIEL STARCH · JAMES B. STROUD · DEWEY B. STUIT · DONALD E.
SUPER · ST. CLAIR A. SWITZER · PERCIVAL M. SYMONDS
GILBERT E. TEAL · LORIN A. THOMPSON · JOSEPH TIFFIN · MILES A.
TINKER · HERBERT A. TOOPS · M. R. TRABUE · ARTHUR E.
TRAXLER · CHARLES H. VOELKER · ROBERT J. WHERRY
WALTER H. WILKE · WALLACE H. WULFECK · LLOYD N. YEPSEN

& COMPANY · INC · TORONTO

150.13
FRY

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated 4.2.54
Accs. No. 197

Copyright, 1950, by Douglas H. Fryer and Edwin R. Henry
Printed in the United States of America
ALL RIGHTS RESERVED

PREFACE

This *Handbook* is intended to serve three purposes: (1) to provide a comprehensive survey of applied psychology that differs in point of view from the treatment usually found in textbooks; (2) to provide a detailed account of specialization in professional work that will show how psychologists apply the principles, techniques, and experimental findings of scientific psychology to the solution of everyday practical problems of society; and (3) to provide a description of the organization and administration of professional psychology, its standards and requirements, its relationships with other professions, and its contributions to their practice.

The *Handbook* has been arranged as follows: the first large unit, consisting of Chapters I, II, and III, presents the general applications of psychology to group living and to individual efficiency and adjustment. References recommended by the contributors for further general reading are given at the conclusion of each of the sections. A separate bibliography for the chapters of this first unit will be found at the back of the book ("A Selected and Classified Bibliography: A. Individual and Group Living").

A second large unit of the *Handbook* consists of Chapters IV through VIII, which have to do with the applications of psychology to business and industry. For this unit there are two bibliographies: "B. Personnel Psychology," for Chapters IV and V, which are concerned with selection and training; and "C. Industrial and Business Psychology," for Chapters VI, VII, and VIII, which are concerned with transportation, marketing, and the industrial problems of the worker.

A third unit has to do with the applications of psychology to education (Chapters IX and X); there is a separate bibliography ("D. Educational Psychology") for this unit. The fourth unit, including Chapters XI through XIII, deals with various clinical applications; "E. Clinical Psychology" is the bibliography designed to accompany these chapters.

The last unit of the book, consisting of Chapters XIV through XVIII, is composed of contributions concerned with various professional matters: problems of independent practice, relations with other professions, the contributions of psychology to the professions and the arts, the administration of psychological services, and the definition of professional psychology. The last of the bibliographies, "F. Professional Relations," includes the references for this unit of chapters.

The various bibliographies represent the major contributions of applied and professional psychology as conceived by the contributors, and they provide a list of authors for ready reference. Appreciation is expressed to Miss Esther Gist and Miss Harriet Thompson for performing the major tasks of styling and checking these bibliographies.

The subject index for the *Handbook* was prepared by Mr. James W. Campbell, whose cooperation in performing this work is greatly appreciated by the editors.

Many psychologists have been concerned with the production of the *Handbook*. Some years ago the Board of Editors of the American Association for Applied Psychology outlined the editorial program of the association as follows: (1) the editing of representative research journals for the several large areas of psychological practice, (2) the publication of a professional journal (issued as the *Journal of Consulting Psychology*), (3) the editing of an applied monograph series (issued as *Applied Psychological Monographs*), (4) the preparation of a professional dictionary of psychology, (5) publication of a professional directory (issued as the *Directory of Applied Psychologists*), and (6) the preparation of handbooks of psychological practice. This book is the first in the Handbook Series, and the present editors were assigned the task of its preparation with the understanding that members of the association would cooperate in the undertaking.

The outline of the book was prepared in consultation with more than one hundred applied psychologists who had been invited to become contributors to the *Handbook*. The cooperation received was all that could be asked, and the editors derived genuine pleasure from their work with the authors of the various sections of the book.

Much of the writing and editing of the *Handbook* was done during the early years of the late war, but for various and obvious reasons the book could not be completed and published at that time. In the meantime several original contributors had developed new interests; and various changes, largely expansions, had occurred in the profession. Many contributions had to be revised or rewritten. New sections were added to include important professional developments. The tremendous growth in clinical and industrial psychology, the developments in military application, and the specializations of human engineering required attention.

The heartfelt thanks of the editors are expressed to all who participated in the preparation of the *Handbook*, especially to those who retired and to those who joined the group in later years. We regret that it is impossible to name here the many persons, in addition to the present contributors, who helped in the undertaking.

Some contributions will be recognized as the last writings of those whose deaths occurred prior to the publication of the *Handbook:* Edwina A. Cowan, Joseph Jastrow, John G. Jenkins, Lillien J. Martin, and Carl E. Seashore.

New York
January, 1950

DOUGLAS H. FRYER
EDWIN R. HENRY

LANDMARKS OF PROFESSIONAL PSYCHOLOGY

(Prior to 1940)

It has been said that psychology has a long past but a short history. Some of the landmarks in the short history that contributed to the development of applied psychology are presented below. They are the result of a poll of the contributors to the *Handbook* and are limited to the period prior to 1940. It is thought that this list should provide a background against which more recent contributions can profitably be viewed.

Many more landmarks are expected to emerge from recent developments, but time and perspective will be required to establish them. Even now it is clear that the groups of workers organized into research teams during and since the late war, such as the Committee on Classification and Training of Aircraft Pilots, the Committee on Human Resources of the Department of Defense, the Personnel Research Section of the Adjutant General's Office, the Psychological Research Branch in the Office of the Air Surgeon, the Applied Psychology Panel of the National Defense Research Council, the Psychological Unit of the Office of Strategic Services, are certain to be included in the landmarks of this period.

The psychological work described in this book had its roots in the events, principles, and techniques illustrated by these landmarks.

1860	Introduction of psychophysical methods by Gustav Theodor Fechner in his *Elemente*
1877	Invention of the correlation method by Francis Galton
1878	Initial experimental work on fatigue by É. J. Marey in France
1879	Founding of the first laboratory for exclusive study of psychology by Wilhelm Wundt in Germany
1881	First scientific journal in psychology published, *Psychologische Studien,* founded by Wilhelm Wundt
1883	Introduction of psychometrics by Francis Galton in England, in his *Inquiries into Human Faculty*
	Early application of psychology to educational method in *Methods of Teaching History,* by G. Stanley Hall
1884	First psychological laboratory in America founded by G. Stanley Hall at Johns Hopkins University
1885	Invention of millisecond (sigma) measurement by J. McKeen Cattell
1887	Publication of the first psychological journal in the United States, *The American Journal of Psychology,* founded by G. Stanley Hall
1888	Mosso's ergograph invented
1890	Initial psychological tests of individual differences by J. McKeen Cattell
1891	Earliest journal with application of psychology as its purpose, the *Pedagogical Seminary,* founded by G. Stanley Hall
1892	First association of psychologists organized in America, the American Psychological Association, with G. Stanley Hall as president
1893	The beginnings of scientific study of the child with Milicent W. Shinn's *Notes on the Development of the Child*
1895	First objective educational achievement test devised by Rice
1896	Earliest psychological clinic, organized by Lightner Witmer
1897	Bryan and Harter's pioneer study of an occupation: telegrapher
1902	Kraepelin's pioneer study of learning curves
1903	Publication of Binet's *Étude expérimentale de l'intelligence,* the foundation for the Binet-Simon Tests of Intelligence
1904	Factor analysis introduced by Spearman
1905	Lahy's early study of typewriting skill
1906	Psychological Laboratory for the study of the

feebleminded, founded at the Vineland Training School by H. H. Goddard

1907 Appearance of early journals devoted to psychological practice: *The Psychological Clinic,* in the United States, edited by Lightner Witmer; and *Zeitschrift für angewandte Psychologie,* in Germany, edited by Otto Lipman and Wilhelm Stern

1908 Earliest publications on the application of psychology to law, *On the Witness Stand,* by Hugo Münsterberg

First mental-age scale of general intelligence: the Binet-Simon Scale

W. F. Book's reports of investigations of typewriting skill

Lahy's pioneer study on selection of streetcar operators

1909 Publication of the first book on vocational guidance, *Choosing a Vocation,* by Frank Parsons

Goddard's translation of the Binet-Simon Scale

Initial investigation of metabolic effects of motor and mental work, by Benedict and Carpenter

1910 Developments in measurement assembled by Guy M. Whipple in *Manual of Mental and Physical Tests*

Kent-Rosanoff Association Tests

Publication of *Psychology of Advertising,* by Walter Dill Scott

1911 Investigation of motion study by Frank B. Gilbreth

Early applications of psychology to business, in *Influencing Men in Business,* by Walter Dill Scott

1912 H. L. Hollingworth's "Influence of Caffeine on Mental and Motor Efficiency"

National Committee for Mental Hygiene founded

Formulation of quantitative method in *An Introduction to the Theory of Statistics,* by G. U. Yule

Kuhlmann's revision of the Binet-Simon Scale

Formulation of a program of industrial psychology by Hugo Münsterberg, in *Psychology and Industrial Efficiency*

Founding of the American Association for the Study of Feebleminded and publication of its organ, *The Journal of Psychoaesthenics*

1913 Publication of *Educational Psychology,* by E. L. Thorndike, the first inclusive statement of the application of psychology to education

1914 Selection tests developed for telephone operators by H. C. McComas

First vocational guidance bureau in a public school system, founded by H. L. Wooley, in the Cincinnati Public Schools

1915 British Health of Munitions Workers' Committee established (later replaced by Industrial Health Research Board)

Selection tests for telegraphers developed by E. S. Jones

Division of Applied Psychology (later known as Division of Cooperative Research) established by W. V. Bingham at Carnegie Institute of Technology

Walter Dill Scott appointed first professor of applied psychology in a university

Initial scientific study of the interview in the selection of salesmen, by Walter Dill Scott

Selection Tests introduced by H. B. McChesney in Western Union Telegraph Company, New York

1916 Selection of chauffeurs for the German army by W. Moede's tests

Publication of the *Stanford Revision of the Binet-Simon Scale,* by L. M. Terman

Vocational Psychology, published by H. L. Hollingworth

Institute für Arbeitspsychologie established in Berlin

Scott's Man-to-Man Rating Scale introduced in Officer's Training Camps, United States Army

Organization of American Educational Research Association

1917 Tests for the selection of typists and stenographers developed by H. W. Rogers

Organization of the Psychological Committee of the National Research Council, with Yerkes as chairman

Initial study by Marston of blood pressure as diagnostic of deception

Research Bureau of Retail Training established at Carnegie Institute of Technology under the direction of J. B. Miner and W. W. Charters

Committee on Classification of Personnel in the Army, created by Secretary of War Baker, with Walter Dill Scott as director, E. L. Thorndike as chairman, and W. V. Bingham as executive secretary

Laboratory of psychology established in the Saxon Railway Company, Germany

First large-scale use of group intelligence tests, Army Alpha, administered to 2,000,000 soldiers in the United States Army (based in part on Arthur S. Otis' work)

Journal of Applied Psychology founded by G. Stanley Hall

First performance scale of general intelligence devised by Pintner and Paterson

First association of applied psychologists: the American Association of Clinical Psychologists (later the Clinical Section of the American Psychological Association), with Leta S. Hollingworth as president

Portable testing laboratory introduced in a factory by Henry Link

Development of the first psychoneurotic measure by Woodworth and Wells, the Psychoneurotic Inventory

Training school for military psychologists established at Camp Greenleaf, Georgia

1918 Trade Test Division established in the United States Army

First nonlanguage group intelligence test, Army Beta, administered to 100,000 soldiers in the United States Army

1919 Invention of Tests of Musical Talent by Carl E. Seashore

First private business in psychology, "The Scott Company," founded by Walter Dill Scott in Philadelphia

Publication of *The Personnel System in the United States Army,* which reported on the work of the Committee on Classification of Personnel in the Army during World War I

1920 International Congress of Psychotechnology founded at Geneva

First mental hygiene clinic for preschool children at Mt. Zion Hospital, San Francisco, founded by Lillien J. Martin

The Graphic Rating Scale developed by Max Freyd

Earliest Interest Inventory devised by Bruce Moore at Carnegie Institute of Technology

1921 Founding and incorporation of two national institutes of consulting psychology: the National Institute of Industrial Psychology of Great Britain, with Charles S. Myers as director; and the Psychological Corporation, in the United States, with J. McKeen Cattell as president

Personnel Research Federation organized with W. V. Bingham as director

Founding of the New York State Association of Consulting Psychologists, the first state association of psychological practitioners, with David Mitchell as president

Publication of *Psychological Examining in the U. S. Army,* with Yerkes as editor, which told of the work of the military psychologist in World War I

Publication of J. C. Chapman's *Trade Tests,* telling of the Army trade tests

Life Insurance Sales Research Bureau founded in New York

1922 Laboratory of Industrial Psychology established in Moscow, U. S. S. R.

Psychotechnical Institute established in Zurich, Switzerland

Child guidance clinics established by the National Committee for Mental Hygiene and the Commonwealth Fund

Harvard Growth Study initiated

1923 Bureau of Sales Research established by H. G. Kenagy in the sales department of Proctor & Gamble Company

State-wide educational testing service developed by H. A. Toops at Ohio State University

Bureau of Public Personnel Administration established in Washington, D. C., by L. L. Thurstone

1924 The first government psychologist, L. J. O'Rourke, employed by the United States Civil Service Commission as director of research

Social Science Research Council founded

1928 Publication of the Vocational Interest Blank by E. K. Strong, Jr.

Attitude tests developed by L. L. Thurstone

1929 Cooperative Test Service of the American Council on Education established

First psychological clinic for the rehabilitation of the aged established by Lillien J. Martin in San Francisco: The Old Age Center

1930 Psychological research department organized by R. S. Uhrbrook in the industrial relations division of Proctor & Gamble Company

1931 Occupational Analysis Clinic, Employment Stabilization Research Institute of the University of Minnesota

Inauguration by Starch of the first continuing measurement of advertising readership

1933 Inauguration by The Psychological Corporation of Psychological Barometers, periodic nation-wide polls of public opinion and buying habits

1934 Occupational Research Program established in the United States Department of Labor

Founding of the American Institute of Public Opinion

1935 Psychometric Society founded with L. L. Thurstone as first president

1936 Founding of the Society for Psychological Study of Social Issues

Dartmouth Visual Survey initiated

Public Opinion Polls for forecasting results of national elections popularized independently by Crossley, Gallup, and Roper

1937 Organization of the American Association for Applied Psychology, which amalgamated existing applied associations, with Douglas Fryer as president

Founding of the Institute for Propaganda Analysis

Publication of the first professional organ, *Journal of Consulting Psychology,* with J. P. Symonds as editor

1939 Publication of the *Dictionary of Occupational Titles,* by the United States Department of Labor

CONTENTS

VOLUME ONE

PREFACE v

LANDMARKS OF PROFESSIONAL PSYCHOLOGY vii

CONTRIBUTORS xvii

Chapter I

GROUP LIVING

SEC. PAGE

1. Social Effectiveness and Leadership *Henry C. Link* 3
2. Social Attitudes *H. H. Remmers* 10
3. Morale *Douglas H. Fryer* 16
4. Propaganda *George W. Hartmann* 23
5. Psychological Warfare *Joseph W. Eaton* 31
6. Unemployment *M. R. Trabue* 37
7. Pseudology *Joseph Jastrow* 41

Chapter II

INDIVIDUAL EFFICIENCY

8. Reading Efficiency *William S. Gray* 49
9. Typography and Legibility in Reading *Miles A. Tinker and Donald G. Paterson* 55
10. Illumination *Dean Farnsworth* 60
11. Ventilation *Kinsley R. Smith* 68
12. Fatigue and Its Alleviation *Raymond A. Katzell* 74
13. The Process Conception of Psychological Illness *Matthew N. Chappell* 84
14. Drugs and Smoking *St. Clair A. Switzer* 87

Chapter III

THE INDIVIDUAL ADJUSTING TO HIS ENVIRONMENT

15. Motivation and Self-Direction *Douglas H. Fryer* 93
16. Interest in Work and Play *Donald E. Super and Jack W. Dunlap* 100
17. Aptitude and Intelligence *George K. Bennett* 109
18. Personality and Physique *Otto Klineberg* 116
19. Guided Learning *T. R. McConnell and G. Lester Anderson* 120
20. Vocational and Educational Satisfaction *Robert Hoppock and Donald E. Super* 126

Chapter IV

TECHNIQUES OF PERSONNEL PSYCHOLOGY

21. Job Analysis *Carroll L. Shartle* 135
22. Personnel Interviewing *Bruce V. Moore* 142
23. Rating *Donald G. Paterson* 147
24. Worker Analysis *Beatrice J. Dvorak* 155
25. Job Classification and Evaluation *Carroll L. Shartle* 160

SEC. PAGE
26. Vocational Test Construction *Herbert A. Toops* 163
27. Criteria and Validity *Robert J. Wherry* 170
28. Item Analysis *Robert J. Wherry* 177
29. Personality Tests and Prediction *Edward A. Rundquist* 182
30. Effectiveness of Selection Devices *Marion W. Richardson* 192

Chapter V

FIELDS OF PERSONNEL PSYCHOLOGY

SELECTION AND TRAINING

31. Clerical Personnel *W. Leslie Barnette, Jr.* 197
32. Mechanical Personnel *Thomas W. Harrell* 208
33. Field Salesmen *Marion A. Bills* 212
34. Retail Sales Personnel *Kate Lewis* 215
35. Management's Use of Personnel Research in Training *Roger M. Bellows* 220
36. Federal Civilian Personnel Administration *L. J. O'Rourke* 229
37. Military Personnel Management *E. Donald Sisson* 236
38. Metropolitan Police Personnel *Sadie Myers Shellow* 247
39. Subnormal Workers *Theodora M. Abel and Emily Burr* 251
40. Employment Agencies *Lorin A. Thompson* 256

Chapter VI

INDUSTRIAL PSYCHOLOGY

41. Worker and Manager *Daniel Starch* 263
42. Efficiency in Work *Rexford Hersey* 268
43. Occupational Safety *Gilbert E. Teal* 277
44. Building Illumination *Miles A. Tinker* 283
45. Industrial Ventilation *L. P. Herrington* 288
46. Industrial Morale *Francis F. Bradshaw and Herbert E. Krugman* 297
47. Military Morale Procedures *Francis Keppel* 303
48. Color Vision *Dean Farnsworth* 309
49. Bio-Mechanics *Jack W. Dunlap* 311

Chapter VII

TRANSPORTATION

50. Street and Highway Traffic *Theodore W. Forbes* 325
51. Railroads *W. Fred Cottrell* 336
52. Aviation *John C. Flanagan* 341

Chapter VIII

BUSINESS PSYCHOLOGY

53. Product Testing *Philip G. Corby* 351
54. Copy Testing *Albert D. Freiberg* 355
55. Advertising Media Research *Darrell Blaine Lucas* 364
56. Customer Research *John G. Jenkins* 371
57. Radio and Television Audience Research *Ernest Dichter* 375

VOLUME TWO

Chapter IX

EDUCATIONAL PSYCHOLOGY: TECHNIQUES OF EDUCATIONAL ADJUSTMENT

SEC. PAGE
58. Services of the School Psychologist *Gertrude Hildreth* 383
59. School Classification *Arthur E. Traxler* 389
60. Principles of Child Guidance *Percival M. Symonds* 394
61. Care and Treatment of Language Deficiencies *Morton A. Seidenfeld* 400
62. Exceptional Children *T. Ernest Newland and Lester N. Myer* 404
63. Educational Test Construction *John C. Flanagan* 412
64. Accelerated Higher Education *S. L. Pressey* 420

Chapter X

FIELDS OF EDUCATIONAL PSYCHOLOGY

65. Child Development *John E. Anderson* 427
66. Instruction in Elementary and Secondary Schools *Gordon Hendrickson* 432
67. Instruction in College and University *Carter V. Good* 441
68. Predicting Success in College and University *Edwin R. Henry* 449
69. Professional School Admissions *Horace B. English and George Nagy* 453
70. College Student Personnel Services *William D. Glenn, Jr., and Wallace Gobetz* 460

Chapter XI

CLINICAL PSYCHOLOGY

71. The Psychological Examination *Mary Fite and Arthur L. Benton* 469
72. Therapeutic Techniques *Donald E. Baier* 476
73. Reading Difficulties *James B. Stroud and Dewey B. Stuit* 483
74. Speech Difficulties *Charles H. Voelker* 488
75. Hearing Difficulties *Louis William Max* 499
76. Visual Efficiency *Robert J. Beitel, Jr.* 504

Chapter XII

FIELDS OF CLINICAL PSYCHOLOGY

77. Mental Deficiency *Lloyd N. Yepsen* 513
78. The Mentally Ill *Isabelle Kendig* 518
79. Physical Disease *Morton A. Seidenfeld* 525
80. Community Child Guidance Clinics *Phyllis Blanchard* 531
81. Child Dependency *Marie Skodak* 538

Chapter XIII

PENOLOGY

82. Crime *Milton Metfessel* 549
83. Examination of Offenders *Harold E. Burtt* 555
84. The Courts *Lowell S. Selling* 561

SEC. PAGE
85. Schools for Delinquents *Luton Ackerson* 568
86. Prisons and Penitentiaries *James D. Jackson* 573

Chapter XIV

CONSULTING PSYCHOLOGY

87. The Consulting Psychologist in Business and Industry *Wallace H. Wulfeck* 585
88. General Private Practice *George Lawton* 592
89. Consulting Child Psychologist *Rose G. Anderson* 603
90. The Rehabilitation of the Aged *Lillien J. Martin* 610
91. Psychiatrists and Psychologists in Private Practice *Hyman Meltzer* 614

Chapter XV

THE PROFESSIONS: CONTRIBUTIONS OF APPLIED PSYCHOLOGY

92. Medical Practice with Special Reference to Pediatrics *Grace Rubin-Rabson* 625
93. Engineering *Lillian M. Gilbreth* 630
94. The Lawyer *Steuart Henderson Britt* 634
95. Librarianship *Alice I. Bryan* 638
96. The Social Agency *Edwina A. Cowan* 644
97. The Vocational Counselor *Harry D. Kitson* 647
98. Eugenics *Gladys C. Schwesinger* 653
99. The Physical Educator *Coleman R. Griffith* 658
100. Rehabilitation of the Handicapped *D. H. Dabelstein* 664

Chapter XVI

THE ARTS: CONTRIBUTIONS OF APPLIED PSYCHOLOGY

101. The Arts of Color and Form *Ray Faulkner and Eugene Myers* 675
102. Music *Carl E. Seashore* 681
103. The Theatre *Grant Fairbanks* 687
104. Speakers and Audiences *Walter H. Wilke* 692

Chapter XVII

ADMINISTRATION OF PSYCHOLOGICAL SERVICES

105. Consulting Psychology in Business and Industry *Paul S. Achilles* 703
106. Administration of Psychological Services in Public Schools *Bertha M. Luckey* 707
107. State Administration in Public Schools *Warren W. Coxe* 714
108. Administration of Psychological Services in Colleges and Universities *William D. Glenn, Jr., and Wallace Gobetz* 719
109. Administration of Psychological Services in Institutions for the Mentally Disordered and Mentally Retarded *David Shakow* 726
110. Administration of the General Psychologic Center *Herman de Fremery and Fern McGrath* 731

Chapter XVIII

THE PROFESSIONAL PSYCHOLOGIST

111. Vocational Information *Ralph F. Berdie* 739
112. Clinical Psychology: Guidance and Training *William A. Hunt* 745

SEC. PAGE
113. Vocational Psychology: Guidance and Training *Joseph Tiffin* 749
114. Educational Psychology: Guidance and Training *M. R. Trabue* 753
115. Professional Attitudes *Warren W. Coxe* 758

A SELECTED AND CLASSIFIED BIBLIOGRAPHY IN APPLIED PSYCHOLOGY

A. Individual and Group Living 767
B. Personnel Psychology 777
C. Industrial and Business Psychology 789
D. Educational Psychology 802
E. Clinical Psychology 810
F. Professional Relations within the Profession and with Other Professions 818

INDEX OF CONTRIBUTORS 829

SUBJECT INDEX, by James W. Campbell 831

CONTRIBUTORS

		PAGE
Theodora M. Abel	*Subnormal Workers*	251
Paul S. Achilles	*Consulting Psychology in Business and Industry*	703
Luton Ackerson	*Schools for Delinquents*	568
G. Lester Anderson	*Guided Learning*	120
John E. Anderson	*Child Development*	427
Rose G. Anderson	*Consulting Child Psychologist*	603
Donald E. Baier	*Therapeutic Techniques*	476
W. Leslie Barnette, Jr.	*Clerical Personnel*	197
Robert J. Beitel, Jr.	*Visual Efficiency*	504
Roger M. Bellows	*Management's Use of Personnel Research in Training*	220
George K. Bennett	*Aptitude and Intelligence*	109
Arthur L. Benton	*The Psychological Examination*	469
Ralph F. Berdie	*Vocational Information*	739
Marion A. Bills	*Field Salesmen*	212
Phyllis Blanchard	*Community Child Guidance Clinics*	531
Francis F. Bradshaw	*Industrial Morale*	297
Steuart Henderson Britt	*The Lawyer*	634
Alice I. Bryan	*Librarianship*	638
Emily Burr	*Subnormal Workers*	251
Harold E. Burtt	*Examination of Offenders*	555
James W. Campbell	*Index*	827
Matthew N. Chappell	*The Process Conception of Psychological Illness*	84
Philip G. Corby	*Product Testing*	351
W. Fred Cottrell	*Railroads*	336
Edwina A. Cowan	*The Social Agency*	644
Warren W. Coxe	*State Administration in Public Schools*	714
	Professional Attitudes	758
D. H. Dabelstein	*Rehabilitation of the Handicapped*	664
Herman de Fremery	*Administration of the General Psychological Center*	731
Ernest Dichter	*Radio and Television Audience Research*	375
Jack W. Dunlap	*Interest in Work and Play*	100
	Bio-Mechanics	311
Beatrice J. Dvorak	*Worker Analysis*	155
Joseph W. Eaton	*Psychological Warfare*	31
Horace B. English	*Professional School Admissions*	453
Grant Fairbanks	*The Theatre*	687
Dean Farnsworth	*Illumination*	60
	Color Vision	309
Ray Faulkner	*The Arts of Color and Form*	675
Mary Fite	*The Psychological Examination*	469
John C. Flanagan	*Aviation*	341
	Educational Test Construction	412
Theodore W. Forbes	*Street and Highway Traffic*	325
Albert D. Freiberg	*Copy Testing*	355
Douglas H. Fryer	*Morale*	16
	Motivation and Self-Direction	93

		PAGE
Lillian M. Gilbreth	*Engineering*	630
William D. Glenn, Jr.	*College Student Personnel Services*	460
	Administration of Psychological Services in Colleges and Universities	719
Wallace Gobetz	*College Student Personnel Services*	460
	Administration of Psychological Services in Colleges and Universities	719
Carter V. Good	*Instruction in College and University*	441
William S. Gray	*Reading Efficiency*	49
Coleman R. Griffith	*The Physical Educator*	658
Thomas W. Harrell	*Mechanical Personnel*	208
George W. Hartmann	*Propaganda*	23
Gordon Hendrickson	*Instruction in Elementary and Secondary Schools*	432
Edwin R. Henry	*Predicting Success in College and University*	449
L. P. Herrington	*Industrial Ventilation*	288
Rexford Hersey	*Efficiency in Work*	268
Gertrude Hildreth	*Services of the School Psychologist*	383
Robert Hoppock	*Vocational and Educational Satisfaction*	126
William A. Hunt	*Clinical Psychology: Guidance and Training*	745
James D. Jackson	*Prisons and Penitentiaries*	573
Joseph Jastrow	*Pseudology*	41
John G. Jenkins	*Customer Research*	371
Raymond A. Katzell	*Fatigue and Its Alleviation*	74
Isabelle Kendig	*The Mentally Ill*	518
Francis Keppel	*Military Morale Procedures*	303
Harry D. Kitson	*The Vocational Counselor*	647
Otto Klineberg	*Personality and Physique*	116
Herbert E. Krugman	*Industrial Morale*	297
George Lawton	*General Private Practice*	592
Kate Lewis	*Retail Sales Personnel*	215
Henry C. Link	*Social Effectiveness and Leadership*	3
Darrell Blaine Lucas	*Advertising Media Research*	364
Bertha M. Luckey	*Administration of Psychological Services in Public Schools*	707
Lillien J. Martin	*The Rehabilitation of the Aged*	610
Louis William Max	*Hearing Difficulties*	499
T. R. McConnell	*Guided Learning*	120
Fern McGrath	*Administration of the General Psychologic Center*	731
Hyman Meltzer	*Psychiatrists and Psychologists in Private Practice*	614
Milton Metfessel	*Crime*	549
Bruce V. Moore	*Personnel Interviewing*	142
Lester N. Myer	*Exceptional Children*	404
Eugene Myers	*The Arts of Color and Form*	675
George Nagy	*Professional School Admissions*	453
T. Ernest Newland	*Exceptional Children*	404
L. J. O'Rourke	*Federal Civilian Personnel Administration*	229
Donald G. Paterson	*Typography and Legibility in Reading*	55
S. L. Pressey	*Accelerated Higher Education*	420
H. H. Remmers	*Social Attitudes*	10
Grace Rubin-Rabson	*Medical Practice with Special Reference to Pediatrics*	625
Marion W. Richardson	*Effectiveness of Selection Devices*	192
Edward A. Rundquist	*Personality Tests and Prediction*	182
Gladys C. Schwesinger	*Eugenics*	653

		PAGE
Carl E. Seashore	*Music*	681
Morton A. Seidenfeld	*Care and Treatment of Language Deficiencies*	400
	Physical Disease	525
Lowell S. Selling	*The Courts*	561
David Shakow	*Administration of Psychological Services in Institutions for the Mentally Disordered and Mentally Retarded*	726
Carroll L. Shartle	*Job Analysis*	135
	Job Classification and Evaluation	160
Sadie Myers Shellow	*Metropolitan Police Personnel*	247
E. Donald Sisson	*Military Personnel Management*	236
Marie Skodak	*Child Dependency*	538
Kinsley R. Smith	*Ventilation*	68
Daniel Starch	*Worker and Manager*	263
James B. Stroud	*Reading Difficulties*	483
Dewey B. Stuit	*Reading Difficulties*	483
Donald E. Super	*Interest in Work and Play*	100
	Vocational and Educational Satisfaction	126
St. Clair A. Switzer	*Drugs and Smoking*	87
Percival M. Symonds	*Principles of Child Guidance*	394
Gilbert E. Teal	*Occupational Safety*	277
Lorin A. Thompson	*Employment Agencies*	256
Joseph Tiffin	*Vocational Psychology: Guidance and Training*	749
Miles A. Tinker	*Typography and Legibility in Reading*	55
	Building Illumination	283
Herbert A. Toops	*Vocational Test Construction*	163
M. R. Trabue	*Unemployment*	37
	Educational Psychology: Guidance and Training	753
Arthur E. Traxler	*School Classification*	389
Charles H. Voelker	*Speech Difficulties*	488
Robert J. Wherry	*Criteria and Validity*	170
	Item Analysis	177
Walter H. Wilke	*Speakers and Audiences*	692
Wallace H. Wulfeck	*The Consulting Psychologist in Business and Industry*	585
Lloyd N. Yepsen	*Mental Deficiency*	513

HANDBOOK OF APPLIED PSYCHOLOGY

Volume One

CHAPTER I

GROUP LIVING

1. Social Effectiveness and Leadership	Henry C. Link
2. Social Attitudes	H. H. Remmers
3. Morale	Douglas H. Fryer
4. Propaganda	George W. Hartmann
5. Psychological Warfare	Joseph W. Eaton
6. Unemployment and Psychology	M. R. Trabue
7. Pseudology	Joseph Jastrow

1. SOCIAL EFFECTIVENESS AND LEADERSHIP

HENRY C. LINK, PH.D.

Vice President, Psychological Corporation. Formerly: Director, Psychological Service Center; Industrial Psychologist and Director of Education, Winchester Repeating Arms Co.; Director of Education, U. S. Rubber Co.; Assistant Advertising Manager, Lord & Taylor; Merchandise Manager, Gimbel Bros.; Lecturer, Columbia University; Associate Editor, *Journal of Applied Psychology*. Founder of the Psychological Barometer, a nation-wide urban poll of public opinion and buying habits; Originator of the Double Form Board (Minnesota SR), Three-Inch Cube, Personality Quotient, Oral Directions, and other tests.

Enough experimental work has now been done to permit a partial definition of social effectiveness and leadership, and to justify such conclusions as the following: (1) Social effectiveness consists of certain habits and skills which can be acquired by practice. (2) Leadership is an aspect of social effectiveness. Leaders are not merely born, they are persons who have acquired social effectiveness to an unusual degree. (3) The popular belief that *only a few are born to be leaders, the great majority must be followers,* is a fallacy. Most people can develop skills of leadership if they will. The basis for these conclusions and a set of guides for the development of leadership and social effectiveness will be presented here.

DEFINITION THROUGH TESTS

INTELLIGENCE TESTS AND INTELLECTUAL EFFECTIVENESS

The scientific definition and explanation of human traits have been made possible by the development of tests which measure those traits. For many years philosophers tried in vain to define intelligence. A small library of controversial books on intelligence records their inconclusive speculations. With the advent of the Binet-Simon test, however, the scientific definition of intelligence really began. The formulation of the intelligence quotient by Kuhlmann and Terman carried the definition a step further. Subsequently the development of a great variety of intelligence, mental alertness, and scholastic aptitude and achievement tests helped to define intelligence more fully. It was found that whatever variations or specific aspects of intelligence were measured by these tests, practically all of them measured fundamentally the same thing, namely, the ability to make progress in school or in academic studies. This criterion was the common denominator of intelligence or intellectual effectiveness.

This criterion of intelligence, namely, scholastic progress, is the only one with which the literally hundreds of intelligence and academic achievement tests have shown a consistent and significant correlation. Another facet of this phenomenon is that these tests give consistent and significant intercorrelations. To the extent, then, that these tests measure the same thing, namely, intelligence or scholastic capacity, they naturally measure each other. The multitudinous items which have gone into the construction of these tests reflect the habits and traits which go to make up intelligence. These traits include the habit patterns involved in reading, spelling, arithmetic, writing, and many other fairly specific activities. The various intelligence tests help us to define these traits.

PERSONALITY TESTS AND SOCIAL EFFECTIVENESS

In somewhat similar fashion, the large group of tests, scales, and inventories classified as personality tests help us to define social effectiveness and the habits of which it consists. During the past twenty years, and especially during the decade 1930–1940, there was enormous growth in the field of personality measures. The Thurstone Personality Schedule, the Root Introversion–Extroversion Test, Bell's Adjustment Inventory, the Minnesota Personality Scale, the Bernreuter Personality Inventory, Brown's Personality Inventory for Children, Link's P. Q. or Personality Quotient Test, and the Vineland Social Maturity Scale are among the tests of this type. To the historian this heightened interest in the social aspects of the individual may well seem to parallel the great awakening of social consciousness in the political field during this decade.

The above-mentioned tests undertake to measure a variety of traits such as sociability, extroversion, introversion, self-confidence, inferiority, self-sufficiency, adjustment to the opposite sex, economic self-determination, and emotional stability. All of these traits represent aspects of social effectiveness. Their very names indicate this. But even such subjective traits as emotional stability, inferiority, and self-confidence are intelligible only in terms of social competence or incompetence. The sense of inferiority, for example, is the sense of being inferior to others, whereas self-confidence and emotional stability represent forms of social superiority or at least social poise. Regardless of what particular trait the personality test sets out to measure, it proves to be a measure of one all-pervasive trait, namely, social effectiveness. This comprehensive trait seems to be just as basic and unitary as is intelligence or intellectual effectiveness. But, whereas we can prove that intelligence tests measure intellectual effectiveness as defined by scholastic progress, we cannot so easily prove that personality tests measure social effectiveness. This is because *our civilization has not yet codified the habits of social effectiveness as it has those of school progress.*

We have an elaborate educational system which, by tradition or design, rightly or wrongly, represents our concept and codification of intellectual competence. Our society accepts this educational system at its face value, and from it derives the criterion by which all tests in the intelligence group were validated. But we have no corresponding educational criterion for social effectiveness. Therefore, we have neither a ready-made standard by which to validate tests of social effectiveness, nor a generally accepted curriculum from which to draw the items for such tests. The creators of these tests have been compelled to proceed on the basis of arbitrary and fragmentary assumptions.

THE FREUDIAN INFLUENCE

The great impetus to the study of social effectiveness came unquestionably from the psychoanalytic theories and practices of Freud. From his medical experience with neurotic and psychotic patients Freud concluded that intelligence and education were not nearly so important as social competence. According to Freud, the important thing was not the conscious mind but the unconscious mind. What good was a man's fine intellect and splendid education if some hysteria, delusion, compulsion, or other neurosis prevented him from getting along with his wife, or his children, or his colleagues, or his employer, or his community? Freud's studies were therefore concentrated on the socially maladjusted. His approach to social effectiveness, and that of the psychoanalytic and medical school to this day, is primarily through the avenues of abnormal rather than normal psychology, the study of the socially incompetent rather than the socially competent. Freud's profound but limited concept is strongly reflected in many personality tests, especially the earlier ones, for example, Woodworth's Psychoneurotic Inventory and Thurstone's Personality Schedule. In fact, the chief object of many of these early tests was to detect individuals in need of psychiatric or institutional attention.

SOCIAL ASCENDENCE AND SUBMISSION

Probably the first scale of social effectiveness free from the psychoanalytic influence is the Allport A-S Reaction Study (1928). The items in this scale consist entirely of common social situations in which some normal people act one way, some another. There is no implication of maladjustment or neuroticism. According to this concept, some people dominate others in their social relationships, while some are submissive, bashful, or retiring. This idea was popularly expressed in Scott's pioneer book, *Influencing Men in Business* (1911) and in another early book, by Strong, *The Psychology of Selling Life Insurance* (1922).

The ascendance–submission concept of social effectiveness, crystallized by the A-S Scale, is carried over into the Bernreuter Personality Inventory and a number of other modern personality scales. The use of such scales implies that a person must either dominate others or be dominated by them. This fits in with the popular fallacy that people naturally divide into two groups, leaders and followers. To be sure, test scores show many people at the happy medium, neither ascendant nor submissive. Nevertheless, the implication is that both extremes are essential. We are here confronted by two aspects of psychology, sometimes called pure and applied psychology. One is the impersonal science of people as they are, the other is the science of people as they ought to be. According to the first, many people simply are submissive and many are dominant, and these are the cold facts. The second emphasizes the study of people's capabilities and tries to establish standards or ideals for these capacities. In the light of such tests and experiments, a definition of social effectiveness in terms of those who naturally dominate and those who are dominated, those who are leaders and those who are not, is inadequate.

SOCIAL MATURITY

The Vineland Social Maturity Scale (Doll, 1936) is unique in the extent to which it embodies the concept of social effectiveness as the result of the learning process. Not so much its name as its construction makes this clear. Practically all of its 117 items refer to specific habits and habit patterns grouped according to their complexity into seventeen age periods. All of these habits are of a kind which can be or must be acquired by practice. The more of them the individual acquires, the more socially mature he becomes, no matter what his chronological age. However, the grouping of habits by ages implies that the normal person will acquire the habits characteristic at his age. There is no implication of social dominance or submission, except insofar as the person who has acquired these habits well may be socially more competent than the person who has not learned them.

BEHAVIOR VS. FEELINGS

The Social Maturity Scale, unlike many other personality scales, asks no questions about how the individual feels, whether he is ever troubled with a sense of inferiority or whether his feelings are easily hurt. It defines social maturity in terms of objective or overt behavior and not in terms of ambiguous responses to introspective questions. Through this scale, Doll, the originator, almost seems to say, "If a person will take care of his habits, his feelings will take care of themselves." This is certainly the premise of the P.Q. Test which deliberately omits such subjective traits as emotional stability and sense of inferiority and consists almost wholly of objective habit items.

In sharp contrast are the Humm-Wadsworth Temperament Scale (1934) and the Minnesota Multiphasic Personality Inventory (1940), both of which strongly reflect the psychiatric or medical approach to personality. The latter is described by its authors (1943) as intended for "the clinical worker who wishes to assay those traits that are commonly characteristic of disabling psychological abnormality." The inventory consists of 555 questions comprising scales for paranoia, hysteria, psychasthenia, hypochondriasis, schizophrenia, hypomania, and several others. Almost all of these questions deal with the bad consequences of social incompetence such as fears, worries, and frustrations.

A person's feelings or emotions are important. However, the trend in psychological theory, the development of tests, and clinical work, justify the conclusion that feelings of social inadequacy leading to emotional instability are not the cause of social impotence. They are the result of failure to acquire adequate social skills. Therefore, while the emotions of frustration may give us broad clues to the problems of social incompetence, social competence must be defined in terms of the habits and skills which make it possible.

The Rorschach Test (1921) for personality diagnosis is probably farthest removed from the types of tests discussed thus far. It consists of a series of ink blots for which the subject is asked to give his most immediate association. The tester must then interpret these associations in terms of personality characteristics or symptoms. The test itself provides no specific inventory either of habits or of feelings. These items must be supplied by the tester according to his own peculiar skills or beliefs.

ITEMS OF SOCIAL EFFECTIVENESS

While Doll's scale is not so sharply divided according to traits as are many of the personality tests, some of its items include a range of habits so complex that they might almost be said to represent a trait. For example, the item "assumes personal responsibility" may well include a network of habits which would justify the trait description of *self-reliance*. In fact, most of the traits which other tests make explicit, plus a good many more, are implicit in the items of the Social Maturity Scale. There is one important exception, and that is the collection of habits which have to do with adjustment to the opposite sex. Nowhere in this scale is the specific and distinctive character of these habits recognized. Yet, as clinical psychologists well know, the absence of certain skills in dealing with members of the opposite sex may seriously affect and even destroy a person's social effectiveness generally.

Clear recognition is given by Doll to the importance of habits which contribute to social effectiveness by enabling a person to be independent of people, to get along without them. In children this consists of being able to dress alone, to eat without help, and to go places alone. In adults it consists of such items as "creating own opportunities" and "systematizing own work." The importance of such habits in terms of social effectiveness is so obvious that it is easily overlooked. For instance, how effective, socially, could a healthy adult be who had never learned to dress himself. A person who is entirely dependent on his friends, like an adult who has never learned how to get along without his mother, is limited in his social effectiveness.

THE PERSONALITY QUOTIENT

The P.Q. Test (Link, 1936) was constructed on the hypothesis that "Personality is measured by the extent to which the individual has acquired habits and skills which interest and serve other people." Personality in this hypothesis is synonymous with social effectiveness. The four traits which this test seeks to measure represent phases of social competence: (1) habits of social initiative, (2) habits of self-determination, (3) habits of economic self-reliance, and (4) habits of adjustment to the opposite sex. These traits are scored separately, but the P.Q. is derived from the over-all score on the test and has been standardized on 3,131 boys and girls from the sixth to the twelfth grades and college freshman in more than a hundred schools or school systems throughout the country (Link, 1938).

The title, P.Q. Test, was chosen to emphasize the similarity as well as the contrast with the I.Q. The similarity is that the P.Q., like the I.Q., depends on habits that must be acquired by practice and training, no matter what the individual's innate capacity. The contrast lies in the fact that the habits which contribute to the I.Q. are sharply different from the habits that contribute to the P.Q. A nation-wide study in 48 high schools with 1,138 pupils showed that no correlation existed between the Personality Quotient and scholastic standing (Link, 1937; Roslow, 1940), that is to say, the high scholars were just as likely to have a low P.Q. as a high P.Q., and the poor scholars were just as likely to have a high P.Q. as they were to have a low P.Q.

PERSONALITY VS. INTELLIGENCE

The experimental demonstration of the distinction between personality and intelligence is one of the major discoveries in the history of psychology. Link and Roslow (unpublished) reviewed 58 published experiments in which both personality and intelligence tests were given, and almost without exception, the results of the personality tests show no correlation with the results of the various intelligence tests, or with years of education or any measures of scholastic standing. For example, in a study of ten thousand subjects (Bergen, 1935) there was no correlation between the results of the Bernreuter Personality Inventory on the one hand and those of the Pressey Classification Test, the O'Connor Vocabulary Test, and years of education on the other. As was to be expected, the three tests contrasted with the Bernreuter gave high intercorrelations. The accumulation of such results, together with the fact that the various personality tests often correlate with one another, makes inevitable the conclusion that personality tests measure a unique factor quite distinct from that measured by all intelligence and academic achievement tests.

The discovery of this unique factor, or trait, is of revolutionary significance. Whether we call it personality or social effectiveness, whether we designate it by a P.Q. or an S.E.Q., is of minor importance. The significant fact is that personality or social competence is independent of academic intelligence and academic achievement.

PERSONALITY AS SOCIAL EFFECTIVENESS

Recently there has been a tendency to define personality as *the uniqueness of the person,* or as individuality. This definition is philosophically and etymologically permissible, but from a scientific point of view it is a reversion to mysticism. Just as correct, etymologically, is the definition of personality as the person's social role in the dramas of life, the parts he must learn to play and in which he can achieve success only by practice. The personality tests were developed according to this more specific concept of personality. Personality tests, therefore, do not claim to measure the whole person or his uniqueness, but only those attributes which make him socially effective. In this sense, personality means social effectiveness.

ELEMENTS OF SOCIAL EFFECTIVENESS

Because the elements which constitute social effectiveness have not been codified or established, psychologists have had to make certain assumptions.

MARRIAGE AND SOCIAL EFFECTIVENESS

One of these is that a successful or happy marriage represents one form of social effectiveness. Obtaining a wife or husband in the first place requires a kind of social effectiveness. Retaining them requires it to a still higher degree. Gradually these assumptions are being proved sound or unsound by scientific studies and clinical experience. The extensive studies of happy and unhappy marriages (Terman, 1940) show that happily married couples rate significantly higher on the Bernreuter Personality Inventory than do unhappily married and divorced couples. The highly intelligent, highly educated, on the other hand, are less successful in marriage than the less educated, not so intellectual. Vital statistics show that the higher the education, the lower the rate of marriage. Increasingly it is being demonstrated that those who rate high in the personality tests are more likely to be successful in marriage.

ECONOMIC SOCIAL EFFECTIVENESS

The ability to get a job, hold a job, please one's employer, and achieve promotion have all been assumed to be forms of social effectiveness. A study of unemployed workers during the depression (Lazarsfeld, 1940) showed that jobless men and women who rated high on the Bernreuter Inventory obtained jobs more quickly than did those who rated low. Experimental evidence is accumulating which shows that the kinds of habits and skills required for economic effectiveness are the kinds that are measured by the personality scales and are not detected by the intelligence tests or measures of academic achievement.

LEADERSHIP

Leadership is the most conspicuous manifestation of social effectiveness. Leadership implies that a person's social competence is such that he has achieved a following. Therefore, a nation-wide experiment to validate the P.Q. Test (Link 1937, Roslow 1940) was designed to use leadership as the basic criterion. Leaders, for the purpose of this experiment, were high school students who had been elected by their classmates to positions of leadership, not leaders appointed by teachers or school officials. They were the class officers, officers of the General Organization and other school societies, editors and managers of the school papers, captains and managers of athletic teams, school band, orchestra, glee club, and other groups. The nonleaders were those who held none of these positions of leadership. The average scholarship of these two groups was about the same. However, the leader group was found to have a significantly higher P.Q. than the nonleader group. Therefore, if leaders are leaders because they are high in social effectiveness, then the P.Q. Test, in measuring social effectiveness, also measures leadership.

ACHIEVEMENT OF SOCIAL EFFECTIVENESS

The present concept of social effectiveness goes far beyond that of social maturity in terms of existing norms of social behavior. It introduces the question: What habits and skills contribute more than others toward social competence and leadership? This question emphasizes the great vacuum in our present civilization already mentioned. Not only do we lack a comprehensive system of social values, but we lack a systematic codification of the basic habits and skills which lead to social competence. The weaknesses of our existing educational system, which concentrates on the habits of intellectual effectiveness, are due largely to this fact. One recognition of this condition is the development of college courses on marriage. Until recently marriage was not considered a subject warranting serious education. Indeed, happy marriages were believed to be made in Heaven, or to be due to some sort of affinity or sexual compatibility. The initiation of college courses on marriage seems to establish success in marriage as a desirable and teachable kind of social behavior. However, personality studies indicate that the basic habits and skills contributing to success in marriage are probably acquired in early childhood, even before adolescence. These are the years in which the necessary social habits should be taught. Later, college courses are likely to substitute theories for skills and self-conscious knowledge for unconscious habit patterns. Here, as in so many other fields, the gap between intellectual effectiveness and social competence remains.

The personality inventories are laying a foundation for the filling of this gap. For example, much of the experimental work with the P.Q. Test has concerned itself with evaluations of the habits and skills which constitute its items. From this and other investigations the following seven guides for the development of leadership and social effectiveness have been deduced:

SOCIAL CONTACT

Social effectiveness cannot be developed in a social vacuum. One learns the skills of serving others by working and playing with them, not by remaining aloof. Persons who take part in community affairs, or who have been active members of a Y.M. or Y.W.C.A., dramatic club, the Scouts, or other organized groups, were found to have a higher P.Q. than those who had not. The frequent and often sneering rejection of group activities is eloquent proof that individuals can deliberately choose not to develop the skills of leadership. Parents who, instead of letting their children do as they please, influence them to join the proper groups are at least starting their children on the path toward social effectiveness.

ORGANIZED GROUPS

Highly organized and competitive groups do more to develop leadership than do loosely organized or casual groups. The habits of cooperation in competition, that is, teamwork and sportsmanship, are the very foundations of social effectiveness and leadership. One question of the P.Q. Test asks, "Which of the following 20 activities, including baseball, tennis, bridge, badminton, and hockey, have you engaged in frequently in season?" Those who had engaged in six or seven usually had a higher P.Q. than those who had participated in only one or two. The next question asks, "In which of these did you practice hard and regularly so that you would become good enough to enter a contest?" Such practice, it was found, contributes even more toward a high P.Q. than does casual participation. The greatest single influence in the development of leadership among young people was intensive participation in competitive games and sports, however; not in one but in several.

Prominent in this list are orchestras, glee clubs, and choirs. Few activities require more intensive practice and teamwork than an orchestra. Not only must the individual learn to play his own instrument adequately, he must learn to play it in harmony with from twenty to thirty others. One sour note may destroy the entire effect. The best violin becomes the first violin, a leader, and every competent player may become the leader at times with a solo.

PRIVATE PRACTICE

The practice of certain social skills in private is essential to the development of leadership. The person who spends day after day practicing the piano in solitude is developing a skill which may transform self distrust into social confidence. Almost any skill which one develops to the point of competence may give him or her a sense of confidence and a degree of leadership. Nor does it have to be a skill in music, art, or literature. The more lowly arts, such as cooking, knitting, sewing, telling stories, gardening, and being handy with tools, if developed to the point of superiority, can all contribute toward social effectiveness, even to leadership. Nearly every person can acquire some skill to the point where he excels in the ability to interest and serve others.

ECONOMIC INDEPENDENCE

The development of economic independence is indispensable to leadership. The habits of doing socially useful work are of basic importance. Aside from the experimental and clinical evidences for this conclusion, it seems only natural that the person who leads others and upon whom others depend should be independent and self-supporting himself.

OPPOSITE SEX

Social effectiveness includes habits of dealing with the opposite sex. Psychological practice shows that there are certain individuals who live in only one half of the social world, either the world of men or the world of women. They have developed emotional barriers which keep them from dealing effectively with members of the opposite sex, even in the most simple social situations. Such a barrier is likely to exist whenever an individual has failed to develop, through practice, some of the elementary habits of playing or working with members of the opposite sex. Many of these habits, like social dancing, are obviously of a very specific sort. However, even walking with, talking with, and telephoning to members of the opposite sex are found to be specific habits which must be and can be cultivated by themselves. Specific habits of interesting and serving members of the opposite sex are sharply defined elements in social effectiveness.

BODILY MOVEMENT

Bodily movement is the common denominator of the activities which develop social effectiveness. If one examines the activities mentioned above, or the items which make up the Social Maturity Scale and P.Q. Test, it will be seen that all of them, whether it be playing the piano alone, playing in an orchestra, working for pay, or playing in a baseball game, involve physical movement. The very word, leader, implies movement and action. The leader is one who moves in advance of his followers, one who must make unusual exertions in their behalf. In terms of the definition of social effectiveness he has won his position of leadership by his ability to interest and serve others. Insofar as he has become the servant of all he may become the master of all.

The common denominator of those pursuits which hinder the development of leadership is the absence of bodily movement, the absence of creative exertion. It includes those time consuming pursuits in which the individual merely sits and takes things in—the movies, the radio, casual reading, and just plain sitting.

TEAMWORK TO LEADERSHIP

Leaders must learn to be good followers before they can become leaders. The condition of becoming a leader is the acquisition of skills which make one a good cooperator or a good team mate. For every situation in which a person acquires these skills to the point of leadership, there are many more situations in which his social effectiveness depends on the skills of being a good follower. A study of leaders and nonleaders (Smith and Nystrom, 1937), in which 40 outstanding leaders and 40 nonleaders were selected in each of three high schools, showed that the leaders participated in an average of 6.8 activities compared with 1.7 activities for the nonleaders. The most frequent extracurricular activities among leaders were competitive athletics. Moreover, the leaders had more than twice as much spare time as did the nonleaders in spite of the fact that they engaged in four times as many extracurricular activities. The authors attributed this to the fact that nonleaders either frittered away much time or spent it in ways which did not require concentrated effort; whereas the leaders, through their participation in several highly organized and competitive activities, had learned how to concentrate and how to plan their time.

Especially interesting is the light cast by this study on the oft made remark: "We cannot all be leaders; some of us must be followers." The leaders, it was found, were also the good followers, whereas the nonleaders were the poor followers. A leader might be captain or top man in one or two activities, but merely a member of the team or squad in his other five or six activities. In other words, the leaders were followers in several activities, whereas the nonleaders were followers in only one or none. Thus, conspicuous leadership in one activity was accompanied by having become a well-disciplined follower in several activities. The individuals chosen by their fellows to be leaders were accepted as leaders because they were good followers most of the time.

DEMOCRACY AND SOCIAL EFFECTIVENESS

These findings have unusual significance for the critical problems of citizenship in a democracy. If the habits of teamwork and good sportsmanship are the foundations of social effectiveness and leadership, they are also the foundations of a democracy. Democracy is a team rather than a regiment. It is a form of society in which citizens voluntarily subordinate their freedoms, even their vote, to the rules of the game, to a constitution. Its success depends on the free cooperation and teamwork of all its citizens.

Efficient cooperation on a large scale is impossible without the skills of teamwork acquired in smaller units. Social effectiveness in citizenship depends on habits of action acquired in representative organized groups, where most persons must be good followers most of the time and good leaders some of the time.

Now, psychology makes the discovery that social effectiveness is not synonymous with intellectual effectiveness; that the habits of teamwork are not acquired from books but from group activities; and that our present educational system in emphasizing knowledge and theory fails to teach the basic skills of cooperation. This implies that revolutionary changes in our educational system such as the following five are desirable.

1. For instance, the group project method would be substituted for individual classroom instruction wherever possible. The projects, however, would be of the planned, controlled type rather than the haphazard, do as you please projects so popular in recent years. They would resemble the classic projects of physics, surveying, and chemistry where two or more students work together toward a prescribed goal and according to prescribed rules.

2. A radical shift would be made in the compulsory elements of education. Such subjects as algebra, geometry, English literature beyond the ninth grade and foreign languages would be optional except for certain majors. On the other hand, group athletic sports would be compulsory for all from grade school through college.

3. Extracurricular activities, such as the school paper, orchestra, glee club, dramatics, would become a part of the regular curriculum and every student would be required to participate in some of them.

4. Education in the seventh to the twelfth grades would be focused on vocational preparation so that every high school pupil would graduate with the equipment to earn his own living. Economic self-sufficiency, as demonstrated by a year in a paying job, should be a prerequisite for admission to any college.

5. Secondary schools would be organized as self-supporting, or nearly self-supporting, communities. Every large high school should be a self-contained, self-governing unit, with pupils learning in the process of producing, and acquiring skills in the course of creating or giving service. This plan requires that the high schools be located in the country or with farmlands available, and with pupils living at the school.

Such changes would indeed be revolutionary, but they seem to the author implicit in the findings of psychology in respect to social effectiveness.

READING REFERENCES

Allport, G. W., and F. H. Allport, *The A-S Reaction Study*. Boston: Houghton Mifflin Company, 1928.

Burgess, E. W., and L. S. Cottrell, *Predicting Success or*

Failure in Marriage. New York: Prentice-Hall, Inc., 1939.

Doll, E. A., "Annotated Bibliography on the Vineland Social Maturity Scale." *Journal of Consulting Psychology,* 1940, **4,** 123–132.

Link, H. C., *Manual for the P. Q. or Personality Quotient Test* (rev. ed.). New York: The Psychological Corporation, 1938.

——, *The Rediscovery of Man.* New York: The Macmillan Company, 1938.

——, *The Return to Religion.* New York: The Macmillan Company, 1936.

Roslow, S., "Nation-Wide and Local Validation of the P.Q. or Personality Quotient Test." *Journal of Applied Psychology,* 1940, **24,** 529–539.

Scott, W. D., *Influencing Men in Business* (2nd ed.). New York: The Ronald Press Company, 1916.

Smith, M., and W. C. Nystrom, "A Study of Social Participation and of Leisure Time of Leaders and Non-Leaders." *Journal of Applied Psychology,* 1937, **21,** 251–259.

Terman, L. M., et al., *Psychological Factors in Marital Happiness.* New York: McGraw-Hill Book Company, Inc., 1938.

White, W., *The Psychology of Dealing with People.* New York: The Macmillan Company, 1936.

2. SOCIAL ATTITUDES

H. H. REMMERS, PH.D.

Professor of Education and Psychology and Director of the Division of Educational Reference, Purdue University; Director of the Purdue Opinion Poll for Young People; Vice-President, Richardson, Bellows, Henry and Company, Inc.; Director of Student Personnel Studies, Pharmaceutical Survey of the American Council on Education; Chairman of the State High School Testing Service of Indiana; Editor, Harper & Brothers; Member, Editorial Boards of the *Journal of Educational Psychology* and the *Journal of Experimental Education;* Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology. Formerly Visiting Professor (summer sessions) at State University of Iowa, University of California (Berkeley), State University of Washington, University of British Columbia, and Michigan State College; Instructor at Colorado College; Member of Technical Advisory Committee of the American Youth Commission; Expert Consultant, Personnel Research Section, Adjutant General's Office, War Department; Regional Director, Qualifying Examinations for the Army and Navy College Training Program; Vice-President and Chairman, Section on Education of the American Association for the Advancement of Science; Chairman, Committee on Social Foundations of Education of the Educational Research Association.

The ways of looking at things and persons, forms of readiness, approaching and withdrawing behavior, feelings of rightness and wrongness, and liking or disliking for objects or values differ from emotions though they are related to them. They have been fused in the working concept of *attitude* which may be defined as an effectively toned idea or group of ideas predisposing the organism to action with reference to specific attitude objects. Thomas and Znaniecki (1918), two sociologists, first studied social attitudes and approximated the concept as currently used by social psychologists. Since then psychologists have been increasingly concerned with attitudinal study, for attitudes are theoretically a component of all behavior, overt or covert.

As thus defined the number of identifiable attitudes is the same as the number of things to which the organism can respond, and the concept is coterminous with or closely related to a considerable number of other psychological concepts, such as interests, appreciations, motives, mores, morality, morale, ideals, complexes (in the psychoanalytic sense), values, prejudices, fears, sentiments, loyalties, ideologies, character, and the like. From the point of view of society the system of morals and customs in operation is the social attitudes which constitute the matrix of attitude patterns of the individuals in society. From the point of view of the individual, attitudes constitute the individual's own evaluation of his conduct and desires in relation to the system of social values as he understands them. The functional psychoses thus have their origin in the attitude patterns of the individual.

DEVELOPMENT OF ATTITUDES IN THE INDIVIDUAL

GENETIC DEVELOPMENT

Environmental modifications of the human organism begin with conception. Apart from embryonic development, the changes wrought by environmental contact are manifested from birth on during the early years and even months of life. Here the individual acquires his first attitudes and exhibits his first learned behavior. In brief, he develops personality. Original drives are overlaid by social experiences organized into attitudinal systems.

The newly born child quickly acquires certain feeding habits and comes to regard favorably those who contribute to the gratification of his needs and unfavorably those who frustrate his activities. The baby's attitudes toward his parents will carry over to others. As the child grows in his capacity to respond to those outside his family group, his social attitudes, e.g., cooperativeness, selfishness, dominance, conformity, and the like, will become definite.

As the child acquires "human nature" which will equip him for life in a society of human beings, the area of his relevant experience expands and the attitudinal patterns become incorporated in his personality. Attitudes will be modified, through learning, in accordance with his own goals and drives. This means that he will acquire attitudes like those of his parents, his friends, and the other primary groups (permanent face-to-face groups) of which he is a member. As the individual grows older, secondary groups (non-face-to-face or temporary face-to-face groups) with which he is associated will elicit certain attitudes. In general, the closer the relationship between an individual and others, the greater will be the potency of such relationship in the formation of attitudes.

The growth and development of attitudes involves the integration of numerous specific responses of a similar type. Such attitudes determine the entire adjustment of the individual. If a child is frustrated by his classroom teacher and believes that the teacher is discriminating against him, he will tend to react negatively to the teacher's future criticisms. Accordingly, the term "attitude" is merely a convenient way of referring to the preparedness that exists within the organism for some future activity. It is to be emphasized that such preparedness is neither automatic nor routine, but possesses cognitive and conative aspects differentiating them from habits and reflexes as commonly regarded. Furthermore, attitudes toward inanimate forces, institutions and values, while culturally determined, vary in intensity between individuals and within the same individual at different times.

The baby grows into childhood, youth, and adulthood with attitudes as an important aspect of his learned behavior. Attitudes are evolved from association with his family group, with children in his play and school group, and in general through social-psychological interaction. As an adult, occupation and the responsibilities of citizenship and parenthood are important in attitude development. It is inconceivable that cultural stability, language, and dependability in human affairs could exist without attitudes. Yet the fact that attitudes do change accounts for social innovations, social conflict, and social change.

STABILITY OF ATTITUDES

To what extent can attitudes be modified, to what extent do they remain constant over a period of time, and what factors are related to change of constancy of attitudes? Some of these questions are answered in studies which show that attitudes can be changed and that these changes persist for some time. Buck (1936) suggested that marked changes in student attitudes toward unemployment and socialism were definitely related to changed mores in the population at large and resulted from widespread economic conditions. Thurstone and Peterson (1931) demonstrated that the changed attitudes from seeing a single movie persisted for as much as a year and a half. Knower (1935) measured modification of attitudes toward prohibition by oral argument and printed material, discovering changes in the directions anticipated. Remmers and his students have made similar findings (1935–49). After studying materials on social insurance, capital punishment, conservation measures, and labor unions, high school pupils showed significant shifts in attitudes in the direction planned, and these changes persisted for at least a school year. Attitudes toward the Negro differed significantly after students were given material pointing out unfairness toward and discrimination against the Negro. Elementary school children under a pupil self-government plan were significantly more favorable toward law enforcement than otherwise similar children. Significant shifts toward liberalism occurred in college students during the period 1931–1937 and toward Germans, Jews, Japanese, and Nazis from 1935 to 1947.

Such studies are important in indicating the effect of various types of stimulation in shaping and maintaining attitudes, and point to the tremendous importance of school and community indoctrination in creating beliefs consistent with our aims and values.

MEASUREMENT AND STANDARDIZED MEASURES

Scientific treatment of attitudes requires quantification. In most measurement of attitudes we are really measuring opinions. Opinions, therefore, are expressed attitudes. These may deviate from underlying self-attitudes with or without the awareness of the subject. In the literature they are frequently distinguished as "public" and "private" attitudes.

METHOD

Certain assumptions must be made in order to measure attitudes: that attitudes are measurable, that they vary along a linear continuum and that measurable attitudes are common to the group; that they are held by many people. Limitations of attitude measurements not implicit in these assumptions include the fact that they may be temporary and changeable and subject to rationalization and deception.

While attitudes may be inferred from overt behavior (effort expended for a cause, relative amounts of money spent for goods and services, and the like), most systematic psychological work has concerned itself with opinions expressed or endorsed as indices of attitudes. We shall, therefore, consider the various types of scales that have been developed.

TYPES OF SCALES

Attitudes scales may be classified according to methods of constructing attitude measuring devices such as the interview, a priori "scales," psychophysical scales, sigma scales, master scales, rating scales, behavior scales, and analogous measurements of various sorts. These devices differ markedly in their theoretical and practical adequacy.

A Priori Scales. The crudest method of measuring attitudes is that of the case method, closely followed by ballot counting, such as is exemplified in various public opinion polls. Such polling devices are in effect two-point "scales." For example, in measuring attitude toward capital punishment the relevant item might be "Capital punishment is necessary." The proportions of yes and no votes would then be taken as an index of the existing social attitude for a given population.

A slight refinement is introduced by adding qualifying statements to the main proposition, such as always, sometimes, rarely, never; or strongly agree, agree, undecided, disagree, strongly disagree. Such devices have been criticized on the ground that they are not "scales" in the rational sense of equal units on the scale. A classic example is that of Bogardus' (1925) a priori scale to measure social distance, in which he attempted to measure differences in group mores related to individual willingness to admit members of particular groups to this or that social relationship.

Psychophysical Scales. In 1926 Thurstone applied psychophysical scaling methods to the problem of measuring attitudes. While the rationale of his method is too lengthy to be included here, it consists in essence of arranging a series of opinions relevant to a given attitude object ranging all the way from most favorable to most unfavorable in equally spaced experimentally determined units along a continuum. The average scale value endorsed by a subject thus becomes a measure of his attitude with reference to the attitude object. For the assumptions underlying the method and for the criteria for selecting attitude items the reader must consult the relevant technical literature.

Thurstone himself has pointed out the possible limitations that attitudes may not be on a single continuum but may be discrete, and that the scale values derived from a population of judges is not necessarily applicable to other populations. Hinckley (1939) reported correlations ranging from 0.94 to 0.98 between scale values of items independently scaled by Negroes and whites with reference to attitude toward the Negro. Pintner and Forlano (1937) similarly reported a correlation of 0.98 between two scalings of groups of students known to differ widely when the items of Thurstone's patriotism scale were the consideration.

Sigma Scales. Likert's (1932) modification of the Thurstone methods makes the doubtful assumption that attitudes are distributed normally. On this assumption, he measured attitudes using standard deviation units. Between scale values obtained by this method and that of arbitrarily assigning numerical values from one to five to the various alternative responses, he obtained a correlation of 0.99. He also reported higher coefficients of reliability by this method than by the Thurstone scoring technique. Equating the number of items used, he obtained higher coefficients of correlation for ten different attitude scales and every group measured.

Likert's method has been criticized on the ground that it makes the assumption of normality of distribution and also because scales constructed without the use of judges have correlated only 0.65 to 0.71 with scales using the Thurstone technique of construction. Such findings, it is argued, do not warrant discarding the judging group, especially if measurement of individuals rather than groups is desired. Remmers and Sageser (1941) and Remmers and Ewart (1941) have shown that adding an arbitrary scale of from 3 to 7 points to each opinion item in a scale increases the reliability in accordance with that predicted by the Spearman-Brown prophecy formula (*q.v.*). Guttman (1944) has proposed a rational scheme based on matrix algebra for eliminating items not on the principal continuum for any type of psychological trait. It is relatively simple in application and much less laborious than factor analysis methods.

Master Scales. Another modication of the Thurstone method is that of Remmers' (1934) generalized or master scale method. The essential difference from the Thurstone method is that the opinions which constitute the scale are incomplete sentences without subject, this being supplied at the time of measurement, so that attitudes toward any one of a large group or class of attitude objects can validly be measured on a single scale. Examples of such scales are Scales to Measure Attitudes toward Any: Advertisement, Disciplinary Procedure, Home-making Activities, Play, Practice, Proposed Social Action, Racial or National Group, School Subject, Social Institution, Teacher, and Vocation.

The scaling method used is, like Thurstone's method, based upon the principle that equally often observed differences are equal, with the median scale value of the opinions endorsed becoming the subject's score. Such a series of master scales eliminates much of the laborious process of collecting statements about each attitude to be measured and the extensive labor involved in the scaling. Practically it is useful in that it is immediately available to measure an issue of current interest. For example, when it was announced that President Roosevelt would address the nation via radio on his proposed enlargement of the Supreme Court, the Scale to Measure Attitude toward Any Proposed Social Action was used to measure a group of high school pupils before and after they listened to the speech to measure the change in attitude which occurred as the result of the speech. A number of studies have shown that these scales are internally consistent and give substantially the same results as do Thurstone's Scales for the same attitude object (Grice, Kelley, 1934).

Behavior Scales. Rosander (1937) developed a scale to distinguish verbalized opinions from behavior. He attempted to measure attitudes toward the Negro by the subject's acceptance or rejection of items such as the following.

1. In the community where you live a Negro marries a white girl. You do nothing about it. (6.2).
2. You are bathing at a beach. Some Negroes approach and enter the water near you. You start to fight with them. (1.2).
3. A Negro family moves into a residential district where you live. You invite them to your house. (11.5).

The parenthetical scale values following the statements were obtained by the usual psychophysical scaling methods based upon the equally often observed difference principal.

Eckert and Wilson (1930) constructed a scale on a somewhat similar assumption, namely, that verbal descriptions of specific situations are likely to get closer to the "real" attitude than such scales as those of Thurstone, Likert, Remmers, and others. Their scale entitled "What Would You Do, a Survey of Student Opinion" was not scaled according to the refined methods previously described. It presented specific situations designed to afford some measure of the allegiance of pupils to the "basic tenets of American democracy." The face validity of such scales appears to be good. Their validity as related to behavior must, as is the case for other types of scales, be validated by further research.

Analogous Measurements. Attempts to measure attitudes by the use of scales not directly drawn up for this purpose have not been wanting. These include personality scales, interest blanks, advertising experiments (consumer surveys), polls of public opinion, observation of overt behavior (as group membership, ways of spending money), acceptance-rejection scales (Fitz-Simmons), and the like.

The possibilities of ambiguity and misinterpretation, either conscious or unconscious, have long been stumbling blocks in the use of verbal scales, because of the possible lack of correspondence between expressed attitudes and other, non-verbal behavior. It should be pointed out in this connection that for a large variety of attitude objects verbal response is the only possibility for most individuals. For example, one's response to the notion of a world court or a super-national government could hardly take other than verbal form. So it is with many other issues.

May and Hartshorne's studies (1925–27) of character represent one attempt to measure the difference between expressed attitudes and behavior. A similar attempt is the study of attitudes by projective techniques, in which the attitudes are inferred through interpretations of various types of stimuli or behavior situations. The term "projective technique" implies a field of unstructured material on which the individual can project his personal way of seeing life, his meanings, significances, patterns, or feelings. Such methods differ from standardized tests in that they do not so much seek to interpret individuals in terms of cultural norms, but rather to describe the specific dynamic organization within a particular individual. A further difference from standardized tests is the avoidance of explicit formulation of the issues involved; in this way, the aim is to eliminate conscious falsification or inability (emotionally or intellectually) to respond. A criticism, of course, is the difficulty in applying statistical measures of reliability and validity.

Probably the best known type of projective technique is that of play therapy, in which the child is allowed to express his conflicts, guilt feelings, aggression or other motivations or attitudes through spontaneous creations of clay, finger-paints, dough, or mud. In this way, the child is said to project his feelings into the formless materials, meaningless in themselves. In the same way, he may use dolls or small toys to represent emotionalized persons or situations and thus work through emotional stresses, which could be verbalized only with prolonged and strenuous effort. For example, sibling rivalry and feelings of rejection or inferiority can often be brought to light in the maladjusted child who could not verbalize them directly, even if he were willing to cooperate.

In the same manner, pictures, music, ink-blots have been presented for individual interpretation. The Rorschach Test (1921), a series of ten cards, was devised to allow free-association, projection of unconscious patterns of affectivity, perception, and thinking upon unstructured stimuli. Murray's Thematic Apperception Test (1935), while still allowing individual interpretation, is less plastic, in that it is composed of definite pictures which the individual interprets according to his own experiences and attitudes. In a test designed for use with children, one of the pictures shows a boy fishing near a "no fishing allowed" sign. Responses of necessity involve the child's attitudes toward law enforcement, independence, fear of being caught, and others.

Numerous studies, such as Marston's (1925) toy situation to measure introversion-extroversion, Horowitz' (1936) measurements of social and race attitudes by preference for pictures, Fites' dramatic play with dolls to indicate the child's attitudes toward aggression, are suggestions of the possibilities of these techniques.

Another technique is that of psychodrama, wherein the individual expresses his feelings and attitudes by acting them out. Moreno (1934) cites examples of a ten-year-old boy who hit his mother before bedtime and at parties, of a wife suffering from violent fits of rage; both behavior patterns were eliminated after dramatic situations allowed complete expression of them.

Such methods, still largely in the experimental stage and quite different from the usual test or scale, provide a new and interesting approach to attitude study and so far as evaluated seem to offer definite contributions to the field. It is to be noted that all behavior is in some sense "projective" as it would be interpreted by a clinical psychologist.

ATTITUDES OF CONTEMPORY SOCIETIES AND SOCIAL GROUPINGS

The attitude patterns of all societies can be best understood against the background of the anthropological concept of culture. Man as a biological organism is, with the exception of a few obvious and largely irrelevant differences in color and size, everywhere the same. Everywhere he faces the same persistent life problems: to gain subsistence, shelter, and security and to perpetuate the species, to form some sort of group life in which the individual can function in the common affairs of life, to regulate human conduct through the transformation of simple impulsive behavior into behavior patterns consonant with the strivings of the group. This process of socialization or acculturation is perhaps the most distinctively "human" characteristic of the species. Psychologically the values, the mores, the loyalties, the ideals, in short, the attitude patterns thus formed, may be thought of as the dynamics of social behavior.

The great diversity of existing cultures—so great as to appear completely different and unrelated to each other—nevertheless have this common denominator of coming to terms with the conditions of life. From the point of view of the group, socialization is successful when the individual produces no undue tensions and frictions within the group. From the point of view of the individual, success is a matter of achieving individual goals in relation to the multiplicity of institutionalized attitude patterns. Illustrations of the failure of socialization in our culture from either point of view are such things as crime and delinquency, divorce, inability to hold a job and last, but not least, the functional psychoses.

Within each culture, however defined, are many subcultures—social groups (often pressure groups) with which in various relations the individual identifies himself. Much of social tension is explicable in terms of clashes and conflicts of such subcultures in terms of special attitudes toward religion, race, labor, capital, government, vocations, industry, and business. These may or may not parallel defined social values and needs of the majority. Psychologically the clash in culture known as World War II has its explanation in the conflict of ideologies, systems of values, basically related to the persistent tasks of life previously mentioned.

It is also the characteristic of each culture to magnify its virtues and to overlook its deficiencies and blind spots. The Germans were by no means the first to think of themselves as a *Herrenvolk*, or master people. Such an attitude as an unstated major premise is present in all cultures. In terms of written records the notion of a "chosen people" is several thousand years old. The names of various American Indian tribes were, according to cultural anthropologists, given them by other tribes. Each tribe for itself called itself *the people*.

Within our own nation the subcultures are as many and varied as are the social groupings that we can identify. Labor and capital, white and colored, urban and rural, town and gown, the social set and "the people across the tracks" are a few commonplace illustrations of the many dichotomies and attitudinal oppositions that exist. Social activities inevitably become institutionalized and develop integrated and involved social hierarchies within the activity. Roethlisberger (1941) shows that this is true in industrial plants. Anyone at all familiar with college and university life can easily see such patterns—the extent to which individuals and the groups to which they belong "rate."

VALUE OF ATTITUDE MEASUREMENT TO SOCIETY

The realization is rapidly growing that attitudes, the way individuals and groups feel about the various aspects of their world are probably more determinative of behavior than mere cognitive understanding of this world. When this is granted, the importance and value of attitude measurement becomes at once obvious. On the basis of experimental evidence psychologists agree on the majority of psychological issues of importance in the determination of happy, need-gratifying individual existence. Areas in which there is substantial agreement include mental hygiene, intelligence, race differences, nutrition, specific aptitudes, and educational methods. Turning from the individual to society, from which he is inseparable, it is essential that the body of knowledge on which there is professional agreement and for which there is objective substantiation be utilized for the realization of the individual's capacities and happiness and for their integration in terms of the general welfare.

Among the prerequisites for the fulfillment of the promise of psychology and the other social sciences are the cultivation of attitudes favorable to: (1) social change and social invention, (2) public responsibility and open discussion of public issues, and (3) the fruitfulness of free inquiry. Nothing less in the areas of social attitudes will permit the extension of knowledge for free men in a democratic society. A few specific means of fostering these attitudes suggest the necessity of a continuous and vast program of adult education, equal opportunity of expression, growth of the free press, the expansion of public service by such organizations as public opinion polling agencies, municipal reference bureaus, student opinion surveys, the Educational Policies Commission, the Town Meeting of the Air, and the modification of the programs of our schools to secure the desired affective outcomes.

Psychologists and other social scientists have a definite responsibility of transmitting their findings to the people and assisting a popular understanding of the importance and implications of these findings. This obligation is no less real, for example, than that of the medical profession. In recent years there is growing evidence that psychologists are becoming acutely aware of this responsibility and are bringing their knowledge and method to bear on crucial social problems.

READING REFERENCES

Allport, G. W., "Attitudes," Chap. 17 in Carl Murchison (ed.), *Handbook of Social Psychology*. Worcester: Clark University Press, 1935.

Cantril, H., *Gauging Public Opinion*. Princeton: Princeton University Press, 1944.

Doob, Leonard W., *Propaganda; Its Psychology and*

Technique. New York: Henry Holt and Company, Inc., 1935.

Friedman, Bertha B., "Foundations of the Measurement of Values: The Methodology of Location and Qualification." Teachers College, Columbia University, *Contributions to Education*, No. 914, 1946.

Guttman, L., "The Cornell Technique for Scale and Intensity Analysis." *Educational and Psychological Measurement*, 1947, **7**, 247–281.

Horowitz, R., and L. B. Murphy, "Projective Methods in the Psychological Study of Children." *Journal of Experimental Education*, 1938, **7**, 133–140.

Nelson, E., "Attitudes." *Journal of Genetic Psychology*, 1939, **21**, 367–436.

Thurstone, L. L., and E. J. Chave, *The Measurement of Attitude*. Chicago: University of Chicago Press, 1929.

3. MORALE

DOUGLAS H. FRYER, PH.D.

Associate Professor of Psychology, New York University, and Secretary, Richardson, Bellows, Henry and Company, Inc.; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology. Formerly: Chairman, Department of Psychology, University College, New York University; Expert Consultant to the Secretary of War; Collaborator, US Forest Service; Chief, Recruiting, Selection and Classification Research Unit, Personnel Research Section, AGO, Department of the Army; Chief, Personnel Research Branch, Civilian Personnel Division, Office of the Secretary of War; Adjutant, Morale Branch, General Staff, US Army; Director, Vocational Department, Central Y.M.C.A., Brooklyn, New York; President, American Association for Applied Psychology and Association of Consulting Psychology.

INTRODUCTION

No definition of morale is satisfactory, but everyone recognizes that it is a phenomenon of group living which vitally concerns the individual's happiness as well as the group's solidarity. Research in war or in peace has never contributed, or probably never will contribute, sufficient facts upon which to base an acceptable theoretical definition of morale. The reason for this is that morale implies about everything psychological except, possibly, intelligence, skill, proficiency, and other abilities. Even these endowments are considered to be influenced by morale. Morale is thought of in many different ways: it may be a special concept such as fighting spirit or worker satisfactions; it may apply to the individual or the group such as the attitudes of Republicans toward the New Deal; it may have class or national implications such as the "Main Line" or National Socialism; and it may refer to the individual's well-being.

The phrase "individual well-being in social living" defines best the word, morale. The highest expression of morale is a life lived according to the best standards of well-being of a group, a nation, or a society of nations. Low morale is any deviation in living from these standards of well-being. Morale relates to what the individual and group know of living conditions in the past, in other groups, or elsewhere in the world; it relates to whatever history tells the individual or what his leaders allow him to know.

The individual learns morale as he learns everything else. Each social group to which an individual belongs has a morale to be learned; it has certain frames of reference and of beliefs and opinions, into which the attitudes of the members fit. The individual learns the group's morale as he learns a skill; and often the morale and the skill are learned at the same time. The priest learns his form of worship along with his beliefs about it. The surgeon learns his skill with the knife when he learns what is and what is not humane and ethical. The craftsman learns his trade along with the craft or union rules for its use.

A development of a group's morale standards accompanies the individual's learning of them. Just as the individual in learning a skill may contribute to the development of a skillful technique, so in learning a morale, the individual may contribute to its development. Actually, the individual is learning the morale of many and various groups. There is an interplay of ideas in him when he learns these various morales, and he contributes to one group what he learns from another. Thus, folkways and mores, totems and taboos, customs and laws are ever-developing and changing. Nations, clubs, religions, gangs rise, become active, and then die out in this continuous round of individual-group interplay of ideas, opinions, attitudes, and beliefs.

In summary, then, morale is individual well-being within social living, in which the standards of morale are the psychological frames of reference of social groups. These are in continual evolution due to the learning of many morales by the individual and the interplay within the individual of the frames of reference of various groups.

DETERMINERS OF MORALE

Everything that affects our life, as individuals and as groups, are determiners of morale, e.g., biological adaptation, psychological adjustment, economic status, social integration, national controls and safeguards, education, marital status, age, disease, occupational and artistic skills, ethics and morals, social security, competition, leadership, sex, and so on. All together, these conditions make up our existence. Our well-being and zest for that existence constitutes morale.

The real determiners of morale, however, are the psychological aspects of these biological, economic, or social conditions of living. There may be either high morale or low morale under any of them. The social group establishes the standards. If present living conditions are lower than the attitudes that the people dictate as satisfactory, the level can be made acceptable by means of a different frame of reference or viewpoint. This was done by Nazi Germany when substituting "guns for butter." Propaganda, and psychological warfare are the great manipulators of the standards of well-being, although force and military occupation have played a part under war conditions. But the permanent institutions of a society, such as government, education and religion, are the normal educational channels for the development of morale standards for individual or group.

GENERAL FACTORS

It is impossible to generalize with exactness upon the factors involved in the determination of morale, while research has indicated that certain areas of our life are more or less likely to be influential in determining the level attained by individual and group morale. Morale seems to be tied definitely to the particular social situation in which the individual lives, which appears natural when it is recalled how the individual learns his morales from various social groups. For example, the factors important in morale differ according to whether the social situation is work in industry, study in college, war conditions of bombing, Army life, or peacetime consumer buying.

FACTORS IN MORALE MEASUREMENT AND TRAINING

It is customary to establish as exactly as possible just what are the psychological factors with which one must be concerned before planning training or measurement of morale in a given social situation. This may be done purely upon theoretical grounds or with evidence from research including the results of factor analyses. The Morale Services Division, Army Service Forces, of the United States Army during World War II (Morale Services Division, 1943) selected the following seven factors in Army morale for the training of the soldier and for the measurement of his level of morale: (*a*) faith in the cause and the future, (*b*) pride and confidence in the outfit, (*c*) belief in the mission, (*d*) confidence in training and equipment, (*e*) realistic appraisal of the job ahead, (*f*) satisfaction with the job assignment, and (*g*) belief in the Army's concern for an individual's welfare. Influences were brought to bear on these psychological factors which would develop zeal, discipline, self-confidence, and satisfaction for life as a soldier which, combined with technical proficiency and physical fitness, were expected to produce the efficient soldier.

The Office of Public Opinion Research at Princeton University (Harding, 1944) similarly outlined at the beginning of World War II the factors expected to be influential in civilian war morale and submitted them to measurement, factor analysis, and validation with relation to a criterion of participation in war activities. The three factors of civilian war morale and their essential components were, in order of importance, as follows:

1. Reasoned determination to achieve the objective:
    - (a) Agreement with the objective
    - (b) Determination to achieve the objective
    - (c) Belief that the war objective is obtainable
    - (d) Awareness of objective
    - (e) Hatred of the enemy
    - (f) Awareness of the magnitude of the task

2. Confidence in leaders:
    - (a) Confidence in President
    - (b) Confidence in subleaders
    - (c) Confidence in news
    - (d) Sense of unity
    - (e) Belief that the peace objective is obtainable

3. Satisfaction with traditional values
    - (a) Basic values
    - (b) Feeling of participation

FACTORS EVIDENT UNDER STRESS

How civilian organization and production has been maintained under severe living conditions gives insight into the factors determining morale. These have been outlined from studies of attitudes and behavior of adults and children in England during the war (British Psychological Society, 1941; Vernon, 1941). The remedial measures taken to maintain morale are given as follows:

a. Belief in bodily security, where individuals were killed and their existence was threatened by bombing, was maintained by training in routine activities, even very trivial ones, of work and living.
b. Belief in economic security, where facilities of living were disrupted, was maintained by rationing systems, feeding stations, special transportation and the like.
c. Belief in social security, where homes and communities were destroyed, was maintained by retaining the displaced individuals in homogeneous groups with their normal leadership and activities.
d. Belief in leadership, where psychological warfare, enemy espionage, unsocial acts of civilians, and even abnormal psychological behavior existed, was maintained by government, police, and voluntary controls.

This outline of factors is close to that which might be made for the soldier in actual combat in which mental and physical well-being is maintained through adherence to the rules of bodily, economic, and social security under the direction of competent leadership.

FACTORS EVIDENT IN INDUSTRIAL MORALE SURVEYS

Industry offers the widest area of morale determination within the national scene. There follows a classification of items which were used in industrial morale surveys and which illustrates the factors involved in industrial morale.

Attitudes of common ownership and habits of participation in industrial enterprise, for which illustrative items follow:

Being encouraged to offer suggestions and try out better methods. Being given information about important plans and results which concern individual's work. Being given reasons for changes which are ordered in work. Being invited to offer suggestions when new plans are being considered. Being given to understand completely the general methods which superiors want followed. Knowledge given of other jobs in the organization which individual feels capable of handling and would prefer. Knowledge given of other jobs preferred, even at same pay. Knowledge given of line of promotion. Being permitted to make important decisions in work.

Feeling of individual economic security in the industrial world, for which illustrative items follow:

Assurance of pay increases when deserved. Recognition of fair company records and knowledge of employee's qualifications. Fair dealing on grievances. Prospects of a satisfactory future. Protection from accidents and adequate medical service. Recognition of steady work. Comfortable working conditions. Reasonably surety of holding job as long as good work is done.

Recognition of fair competition, for which illustrative items follow:

Being able to find out if work is improving. Receiving pay comparable to other jobs of equal importance in organization. Receiving pay comparable to that of similar work in other organizations. Receiving recognition and praise for unusually good work. Selection of best qualified employees for promotion when vacancies arise. Bonus plans that offer just reward for more efficient and faster worker. Opportunity for advancement. Certainty of promotion going to best qualified.

Experience of progress toward a goal of skillful accomplishment, for which illustrative items follow:

Receiving necessary help to get results expected of management. Certainty of recognition going to most skillful employees. Regular amount of work. Friendly and helpful criticism of work or criticism of errors. Amount of work required not unreasonable. Freedom to seek help when difficult problems arise in work. Satisfactory daily working hours. Freedom to show initiative on job. Opportunity to use own ideas. Opportunity to learn a job. Chance to use experience.

Feeling of tolerant acceptance and freedom within working group; and experience of mutual give and take in work, for which illustrative items follow:

Dignified treatment when employed. Mutual benefit association membership. Company vacation policy. Pleasant working companions. Freedom to unionize or form protective association.

Confidence in leadership, for which illustrative items follow:

Being given to understand the results expected on the job. Encouragement to seek advice without penalty from superiors. Not being hampered in work by superiors. Not getting conflicting orders. Feeling that superiors understand all of the details and difficulties of one's job. Not being responsible to too many superiors.

FINAL CONSIDERATIONS

The determining factors of morale for a nation in peacetime would be expected to be more varied than under war conditions. Certain factors would be more influential in one occupation than another, at one economic level than another, with one sex than the other, and with one educational level than another. Each social group is a separate morale problem and the morale factors of measurement or training must be established for the individual group. But there will be certain factors most influential upon morale in any particular social group.

They will vary from situation to situation; as the environment is more rigid they will be related more to survival value and be concerned less with the superficial satisfactions of the individual; as the economy is less rigid and more is taken for granted by the individual, minor satisfactions become more important. But all in all, as the above analyses of factors show, the determiners of morale are the attitudes and satisfactions fundamental to well-being and zest in living under whatever conditions that life takes place.

MEASUREMENT OF MORALE

Measures of morale should reveal, according to Wolfle (1942): (*a*) the general status of morale in a social area, such as a community or nation, at a given time; (*b*) the discrepancies between the general state and that of subgroups and individuals in that area; (*c*) the causes of good and poor morale in individuals and groups; (*d*) the effects of morale stimulating techniques; and (*e*) they should serve as a technique in themselves counteracting morale destroying influences. This is an ambitious purpose for morale measurement and training. It presupposes a degree of commonness of attitudes and satisfactions among individuals and groups, about which we know little, and it presupposes a criterion for evaluating individual and group differences in morale, about which we know less. Morale measures do not, and probably never will, serve well all or any of these purposes, but it is good to check what one does in morale measurement against the theoretical ideal stated above.

MORALE MEASURES

Existing morale measures may be divided roughly into two classifications: (*a*) conviction inventories or surveys, which are concerned with attitudes, beliefs, and opinions, and (*b*) participation measures, which are concerned with the actual behavior of the group or individual. Participation is used, also, as a criterion for the validation of conviction measures.

MORALE CRITERION

That no absolute criterion of morale exists is clear from the foregoing discussion. The group frames of reference for convictions and participation in social activities, or the standards by which one views his position in social organization, have always been arrived at through the conflict of many social forces. Present criteria of morale depend upon socially acceptable goals of morale. They are established by leadership or through group organization and they may be good or bad for the individual. Little thought has been given as yet to criteria of morale (or a standard of well-being) in terms of physical and mental fitness of the individual for social living.

CONVICTION MEASUREMENT

Traditional methods of measurement of human feelings have furnished the basic techniques for morale measurement (Beebe-Center, 1932; Fryer, 1931; Likert, 1932; Thurstone and Chave, 1929). The scales for measuring attitudes, satisfactions, interests, personality, opinions, and prejudices, which were developed during the twenties and thirties, are measures of morale in its various aspects, but individually they are limited in their application to one or a few areas of convictions. Usual problems of morale concern group living in all its manifestations. Probably the most cogent reason why the existing scales of attitudes, interests, personality, etc., have not served a useful purpose in morale measurement is that they are unrelated to practical morale problems; in other words, they either do not deal with attitudes and satisfactions which are manifestations of morale in specific social situations, or, where they are concerned with a practical morale problem, they deal with only a few of many attitudes and satisfactions that are important in measuring morale in the social situation.

EARLY INVENTORIES

The earliest morale inventory was prepared to determine the degree of mental adjustment that drafted soldiers made to Army life. It was a part of the morale program directed by Munson, Frost, and Fryer (Hall, 1920; Munson, 1921) at Camp Greenleaf in 1918 during the first World War. This inventory consisted of items selected "hit or miss" that had to do with problems of leaving home, the Army caste-system, Army social life, sympathy with the goals of the war, desire to get into battle, comforts of living while in the Army, anxiety about the future, in fact, with practically everything that might concern the drafted soldier.

The application of scaling techniques to morale measurement took place in the early thirties with the work of O. M. Hall (1934) and Rundquist and Sletto (1936). These investigators used a vague concept of factors in selecting items. Hall's inventory was constructed of items to determine the effect of the depression upon the attitudes of engineers. He selected items that might measure changes in thinking toward government, employers as a class, labor organizations and benefits, religion, and a variety of social matters. These items were scaled with Likert's method of determining internal consistency, and then selected accordingly and assembled into three scales considered to measure "Attitude toward Religion," "Attitude toward Employers," and "Occupational Morale." It is interesting that religious attitudes were little affected by depression, as was indicated by employed and unemployed engineers, but that on the "Attitude toward Employers" scale 68 per cent of the unemployed engineers were more bitter than the employed, and on the "Occupational Morale" scale 75 per cent of the unemployed were less satisfied with occupational conditions than were the employed.

Rundquist and Sletto extended the study of the effect of the depression upon morale to young people and, profiting from Hall's investigation, outlined six factors expected to be affected in the lives of young people by the depression: (*a*) general morale, (*b*) inferiority, (*c*) family relationships, (*d*) respect for law, (*e*) economic conservatism, and (*f*) attitude toward the value of education. Items were written to sample attitudes, experiences, and satisfactions, according to these six factors; and six scales were constructed to measure the six factors according to Likert's method for selecting items. Among these six scales, the economic and general morale scales showed most striking differences between youthful employed and unemployed men. The remainder of the scales showed little or no differences, according to this criterion of depression conditions.

This early experience with morale measurement has led to a clarification of the issues. Morale can be measured only in a particular social setting or situation. Different factors will be involved in the morale of each social situation, although morale factors probably have some general application in many social situations, as indicated in the previous section. But items for morale scales will have little exact transfer from situation to situation, and must be prepared independently for each new situation. This has been the experience in morale measurement of the present decade, both in industry and in the Armed Forces.

If the population with which one is concerned in the measurement of morale is a fairly homogeneous group, it may be found that highly significant items can be written and constructed into a brief scale for the measurement of morale. Graves (1941) in the US Forest Service, was able to get twenty-five items, all of which separated high-morale and low-morale criterion groups with critical ratios for differences in response running from 3 to 13. This was possible because the items could be specific to the social situation as it affected the lives of all members of the homogeneous group.

PARTICIPATION MEASUREMENT

It is often said that what is important is not what people think (their convictions) but what they do (their participation). But to devise a means of measuring what an individual or group of individuals does in all the specific activities of a social situation, as we have defined it above, is a problem yet to be solved. There are no tests of morale participation. But specific activities are considered to be symptomatic of morale, and are sometimes referred to as morale indicators.

Incentive experiments have shown the effect of specific stimuli upon performance, e.g., distractors, rhythms, rewards, competition. Incentives are a part of the social situation in which morale is measured and considering all stimuli affecting the individual as incentives, incentive measurement is measurement of morale in a limited sense.

Direct observation of participation has been employed as an indicator of morale in industry. Roethlisberger and Dickson (1934) placed observers in the workroom at the Western Electric Company for this purpose. In this instance the workers knew of the presence of these investigators, but elsewhere observers have been included in the working group unknown to its members.

War has given us many indicators of morale. The number of activities of a program in which the individual participates, or vice versa, often has been used as an indicator of morale for the program. The volunteering of blood has been used for this purpose in various countries. Such behavior as the following has been considered to be indicative of morale: the amount of singing done by Army units, numbers of faintings from innoculations, numbers of symptoms volunteered by recruits on initial medical examination, numbers on AWOL, number of volunteers for a dangerous enterprise, etc.

In industry, absenteeism and turnover have been considered the prime indicators of morale. Number of suggestions, where a suggestion system is in use, indicates participation in the industrial program. Accident rate and amount of spoiled work are used as indicators of morale. The Employee Relations Department taps various aspects of morale participation, such as attendance upon factory outings, eating in the company restaurant, purchase of stock. Numbers of applications for employment and volunteering for new training are useful indicators: In all divisions of the personnel department are activities in which the degree of the individual's participation in the company program may be indicated.

Morale indicators that have been used by politicians, public officials, artists, and others to gauge the support of their public are: the amount of fan mail received, amount of sales of books and magazines, number of anonymous answers to questions published on social issues, number of telegrams received, and the magnitude of applause at a public gathering.

MORALE INTERVIEWING

Morale interviewing has for its purpose both the investigation of morale and the stimulation of its improvement. The interview enables the interviewer, working in industry, education, agriculture, the Army or the Navy, to understand the problems of those with weak convictions and inhibited participation which is not afforded by measurement. Roethlisberger and Dickson (1934) developed an unguided form of interview which was used with considerable success for several years in the Western Electric Company as a means of stimulating morale by allowing the worker complete freedom of discussion in confidence on any problem troubling him. Morale interviewing has the advantage in industry over conviction measurement of closer contact with morale influences, as yet only vaguely defined in the mind of the worker. It is an essential tool in determining the items for construction of a morale inventory when attitudes and satisfactions become well enough defined for purposes of measurement.

RELATED AREAS

The convictions of consumers about commercial products usually are surveyed by personal interviews in which a consumer sample is selected for statistical reliability to represent a larger population. Such a survey may be regarded as a measure of consumer morale for the product or products under investigation, and morale in the total area of commercial products is measured to the degree that consumer convictions are represented in current surveys.

PUBLIC OPINION POLLS

Public opinion polls survey convictions toward public issues in the same manner as consumer surveys poll convictions regarding products. They provide a quicker assessment of public morale than can be attained from participation in public issues, such as elections, referenda, pressure group legislation, and spontaneous communications to government officials. Such polls are particularly successful in measuring convictions upon specific issues in which the public is vitally concerned. Then they are highly reliable and serve as indicators of morale.

COMMUNITY SURVEYS

Community surveys are distinguished from consumer surveys and opinion polls by the fact that they measure the convictions of a community upon problems of living and are in this sense an inclusive measure of the morale of a social group. They use the methods of the interview, with the usual statistical check on the reliability of the sample selected to represent the community.

Accessioned No. 197

IMPLICATIONS OF RESEARCH

What was known at the end of World War I about the nature of morale, its measurement, and its control, is well summarized by G. S. Hall (1920) and Monson (1921). Measurement did not actually enter the picture until the 1930's and what was learned about morale by the time of World War II has been summarized by Child (1941), Roethlisberger and Dickson (1934), Watson (1942), and Zorbaugh (1941). Methods of measurement and control were developed during World War II to handle many practical problems of military (War Department, 1943) and civilian (Harding, 1944) morale. Unfortunately the research which forms the basis of this work is as yet for the most part unpublished, but it has made available considerable "know-how" in the areas of clinical, industrial, and military psychology.

Numerous researches upon morale indicate that high morale is likely to be present among those of good health, those with superior education, those with efficient habits of work, those of superior economic and social status, and those with a realistic point of view. It is more likely to be present under conditions of good living, but no set of environmental conditions are essential to it. Morale may fluctuate considerably from time to time in the individual within the same environment. Even age seems not to be a general determiner of morale while it affects morale in special situations.

ECONOMIC SECURITY

High morale is apt to be found among those on the higher occupational levels and among those having economic security. This was true among college students whose parents belonged to the higher economic levels. It is true of old people on work relief as contrasted with those in need or dependent upon their children. Employed men stand consistently higher in morale measures than the unemployed. Also, during the depression of the 1930's, the morale of men in desperate straits was higher when they were on work relief.

SOCIAL SECURITY

Morale is higher among those living with their families and with those having satisfactory social contacts in homogeneous groups. The latter statement applies equally among adults and children, among soldiers and civilians, among those living under war conditions and bombing, and those engaged in peaceful pursuits. It is found where there is social confidence, family approval, in marriage, where there is participation in group government, and where individual control and discipline are imposed democratically by the group. It is found among old people who are doing useful work; where young people are living at home; and where people have hobbies and interests. Good companionship is essential to high morale.

LEADERSHIP

Above all else, leadership is a determiner of morale. Economic and social security rest on confidence in leadership. All measures of morale used in the Army during the last war, both among officers and soldiers, were related to the respect held for leadership. A similar situation exists in industry. The character of the boss is most important of all other determiners of morale in industry. Various industrial companies actually differ less in morale measures than departments within one company, showing how the immediate supervisor or the group leader can affect zest and well-being.

SOCIAL VALUE

A social organization exists only because of the morale of its members. A nation, a labor union, a social group, a social class, an army, a gang, a university exists because of its morale. Buildings, orders, authority, rules, laws, ownership, materials, books, etc., are but tools for the expression of well-being and zest for social living. This has long been only vaguely recognized, but during the last twenty-five years its significance has been impressed on all.

The leaders of governments and social groups have turned their attention in recent years to the techniques for influencing morale and they have used these techniques for purposes not always clearly directed toward the physical and mental fitness of all people in social living. The dangers as well as the advantages of morale stimulating techniques to social organization are becoming apparent.

Morale measurement is now available to evaluate social institutions and the changes introduced in them by morale stimulating techniques. Democratic government and management, labor and consumer organizations have need of the knowledge gained from morale measurement. To use this knowledge for the greatest benefit to all cannot but result in the highest permanent morale.

READING REFERENCES

Child, I. L., "Morale: A Bibliographical Review." *Psychological Bulletin*, 1941, **38**, 393–420.

Hall, G. S., *Morale, the Supreme Standard of Life and Conduct.* New York: D. Appleton-Century Co., Inc., 1920.

Hall, O. M., "Attitudes and Unemployment: Comparison of the Opinions and Attitudes of Employed and Unemployed Men." *Archives of Psychology,* 1934, No. 165.

Roethlisberger, F. J., and W. J. Dickson, *Management and the Worker*. Cambridge, Mass.: Harvard University Press, 1939.

Rundquist, E. A., and R. F. Sletto, *Personality in the Depression; A Study in the Measurement of Attitudes.* Minneapolis: University of Minnesota Press, 1936.

Watson, G. (ed.), *Civilian Morale*. New York: Reynal & Hitchcock, 1942.

4. PROPAGANDA

GEORGE W. HARTMANN, PH.D.

Professor of Psychology, Roosevelt College, Chicago. Formerly: Professor of Educational Psychology, Teachers College, Columbia University; Foreign Fellow, Social Science Research Council; Chairman, Society for the Psychological Study of Social Issues; Director of Research, Cambridge-Somerville Youth Study and Ella Lyman Cabot Foundation; Instructor, Dartmouth; Professor, Pennsylvania State College; Lecturer, Harvard; Visiting Professor, Alabama Polytechnic Institute; Editor, *The Social Frontier*.

The psychological aspects of propaganda as a process of social communication appear in (1) the motives of the propagandists, (2) the techniques or special stimuli employed, and (3) the effects of such stimulation upon the recipients' beliefs and overt behavior. The raw or crude efficiency of propaganda is determined by the aims the propagandist holds: if the intended goal (no matter what its nature) is reached, his methods are considered successful by himself and others; if the end is not achieved, the propaganda is considered a failure. This sharp antithesis is defective (1) where it disregards variations in ultimate and hidden costs, monetary and human, (2) when a temporary lost cause prepares the ground for a later success, and (3) where due allowance is not made for the greater inherent difficulties in furthering certain objectives as compared with others. These associated conditions must be included in every full analysis of concrete instances of propaganda.

RELATED SOCIAL FORCES

PROPAGANDA AND EDUCATION

Warren's *Dictionary of Psychology* crisply defines propaganda as "organized publicity used as a means of influencing attitudes." The attitudes involved are invariably *social* attitudes, i.e., mental sets or dispositions toward questions affecting the relations or comparative standing of individuals or groups, which suggests that propaganda is normally much narrower and particularized in the kind of responses to which it directs attention than education, which the same source defines as "the development of abilities, attitudes, or forms of behavior, and the acquisition of knowledge, as a result of teaching or training." The implication is clear that whereas propaganda is generally exhausted in the process of specific attitude formation, the educative process aims at far broader and deeper changes in personalities.

However, the surface techniques of both propaganda and education are so similar in many cases that a distinction between the two processes can then be made only on uncertain ethical criteria. Both seek to change people; they necessarily use the same physical media (press, radio, film, spoken voice); they typically operate by persuasion rather than by violent coercion or through specific inducements, such as money bribes or the direct granting of socially esteemed positions; and they are both conditioned by the readiness of the learners to accept the suggestions conveyed to them. These are important and inevitable *partial resemblances;* as *total patterns* of social stimulation, however, education and propaganda differ significantly. The needs and interests of the persons to be educated are primary in one, whereas the special purpose of the propagandist is the center of reference in the other. Education is for the benefit of the pupils and the entire community; propaganda is for the benefit of the propagandist and those he represents. Ideally, education (when conducted as a scientific and humanitarian enterprise) is devoted to the whole truth and seeks to look at all sides of all complex problems; normally, propaganda operates by sharply biased selection, suppression, emphasis, or distortion of relevant data and by denying a hearing to rival positions.

Propaganda is sometimes organized, systematic deception. Most education probably has some faint propagandistic note, and even the most blatant propaganda conveys some information with educational possibilities. But it is doubtful if the solipsistic view that "your education is my propaganda and my education is your propaganda" is tenable. There are important internal and external differences between the two operations, no matter how blurred and confused the picture often becomes.

Opposed to this view is the commoner opinion that what is labeled education is simply good propaganda, i.e., directed toward goals approved by the judge, while the term propaganda itself is today merely name-calling (itself a primitive propagandistic device!) for whatever the speaker or writer deems to be reprehensible propaganda. Certain critics who adopt this position argue as follows: human motives cannot be objectively appraised—all judgments passed on others' purposes are the extreme of subjectivism. At best, the difference between education and propaganda is a mere quantitative one, for even the most honest education does not succeed in presenting all the facts about a conflict problem; it is always a *weighted* version of the truth that is encountered in actual schooling. Indeed, all good teachers have certain great or small values for which they propagandize. Other critics go even further and declare that propaganda is no more than intensified and emotionalized education, and that the latter would benefit by copying the more vigorous and impressive practices of the former. Conceivably education at its socially worst and propaganda at its socially highest tend to approximate each other; but they usually have different reference points.

The superficial resemblances of propaganda and education have misled many into declaring that since the overt methods often are alike, the intellectual or other psychological processes must be identical. This is like arguing that a book on physics and one on history are the same because both are written in the English language. Actually, even the procedures differ significantly. Education both permits and encourages choice among alternative conclusions after reflection upon the data; propaganda neglects or definitely discourages this. Propaganda ordinarily intensifies social conflict (paradoxically sometimes by denying its existence!); education seeks to resolve these group tensions through study and discussion. The aim of one is understanding of phenomena and group control of events; the aim of the other, power and domination over most people by some people.

At the risk of contrasting ideal or pure processes, one may complete this discussion of differences by declaring that characteristically, education is democratic and propaganda authoritarian in spirit; education seeks to build new discriminations and harder concepts in the learner and propaganda utilizes the older and easier ones in order to convince or motivate its victims; propaganda preferably goes with the tide of current public opinion, whereas education must often go against it in the interest of some long-term or permanent good; propaganda cripples intelligence and operates best on credulity, but education aims to eliminate superstitions and to establish what is called the scientific habit of mind.

PROPAGANDA AND ADVERTISING

Awareness of the over-all difference between education and propaganda may be compared to the naturalist's ability to distinguish between species in cases of zoological mimicry where one creature is often confused with another. On the other hand, the separateness of advertising and propaganda is harder to establish. Both seek to impress an audience favorably with the merits of a given product, service, personality, or cause, and (by implication at least) to disparage competing items; this seems to be true of these kindred processes whether a brand of soap or the government of a great empire is the beneficiary of the effort. Both play frequently and deliberately (although not necessarily) upon the weaknesses, i.e., the ignorance, passion, and irrationality, of the recipients of the stimulation. Characteristically, advertising refers to a limited commercial context of specific vendible articles or personal services, grows out of restricted pecuniary motives, is mainly concerned with controlling a market, and is now commonly regulated by some informal code of ethics; propaganda on the other hand, has a more general political, ideological, and "social conflict" reference, and the ethical restraints—such as appear in words or acts of fairness to one's opponents—appear to be much weaker. The aim of the former is definite selective purchase or patronage by as large a group as can be reached; the aim of the latter is a monopoly of the vaguer loyalties and supporting conduct of a huge population.

The existence of institutional advertising demonstrates that it is impossible to draw a sharp line between propaganda and advertising. Large companies are not only interested in emphasizing the competitive superiority of their product, but also seek to gain the good will of a labor-conscious consuming public by releasing statistics about their wage-scales and employee well-being. No matter what institution is being supported, institutional advertising regularly reaches over into propaganda. The National Association of Manufacturers seems primarily concerned with fighting economic planning by government and with preserving public acceptance of the profit or enterprise system of private capitalism in the face of increasing popular awareness of its many serious limitations. The Communist Party devotes much energy to persuading people that life in the Soviet Union is much better than it really is. Both seek to make or to keep certain mental contents focal in the consciousness of wide masses. While there are minor differences in design and emphasis between advertising and propaganda, there is nothing in them comparable to the fundamental cleavage between propaganda and education.

All propaganda and most advertising that goes beyond simple announcement are indirect and subtle devices for winning approval and wielding social power over others, and are usually lacking in scientific objectivity (although they may use many of the technical findings and research procedures of the laboratory). Despite their concern with controversial situations, they commonly and intentionally fail to provide an equitable hearing for all sides of a case. They are instruments of domination, not of authentic enlightenment. They are highly vulnerable in the light of moral standards emphasizing the organic interdependence of all ends and means as well as the concealed costs and social consequences of their activity.

PROPAGANDA AND MORALE

The practical necessity of morale for social living supplies the motivation for much propaganda. Working efficiency, fighting spirit, willingness to sacrifice, the sense of happiness and well-being, faith in one's crowd, are lowered by unfavorable stimuli (including rival propaganda), and the wheels of counterpropaganda are therefore set in motion to maintain morale.

Where propaganda has an adequate factual basis, or cannot be effectively contradicted, it is likely to be successful in establishing and maintaining morale despite exaggerations and errors of omission. Thus, portraying the enemy as a cringing coward may boost home or civilian morale in wartime, but it is utterly inappropriate among the combat forces which have had immediate knowledge of the foe's valor. Atrocity stories, authentic or imaginary, have the the same function and also the same limitations. When the morale of a group is high, propaganda is superfluous; when it is so low that apathy or even hostility to the propagandist exists, propaganda is all but useless, for none but the exceptional trick will change such a situation. It is in the middle region of the morale scale that propaganda is most serviceable, bolstering sentiment after defeats and energizing what would otherwise be no more than a moderate effort into an all-out offensive.

THESE DISTINCTIONS

Some such imputation of motives as characterizes this discussion must be made before propagandistic activities become intelligible to an observer. It is a postulate of this approach that ultimate or long-run efficiency and happiness in social living are dependent upon full access to the relevant facts and values about a problem. This propaganda in its various forms implicitly denies and actively tries to prevent.

PROPAGANDA TECHNIQUES

ESSENTIAL PATTERN

Since propaganda is systematic special pleading, it uses words, pictures, and other symbols in such a way that the good emblems typical of positive experiences are associated with one cause and the bad labels of negative meanings are attached to its rival. Wartime propaganda makes one nation or alliance saintly and its enemies satanic. The rational view that the brightness spectrum includes many shades of grey is suppressed in favor of the more primitive view that reduces all opposition to extreme black and white. This is the basic propagandist pattern—all others are but variations of this fundamental theme.

A representative case appeared in the motion-picture propaganda against Upton Sinclair's candidacy for Governor of California in the fall of 1934. For reasons best known to themselves, Hollywood officials were generally opposed to Sinclair. Shortly before the November election a presumptive newsreel was released in local theaters purporting to give authentic sound-film field interviews of an inquiring reporter with typical voters. The reporter was shown asking a dear motherly soul knitting on a California porch in attractive surroundings for whom she intended to vote. "Why, of course, Mr. Merriam is my choice." The next scene showed the same reporter prowling about in a dirty Okie-Arkie camp and putting the same question to a tobacco-drooling unappetizing male. "Uppie Sinclair gits my vote, young man. Ain't he promised us guys all a pension if he wins?" Clearly the suggestion was that all the good citizens would support the Republican nominee and that none but the dregs of society would vote for the epic Democrat. Sinclair supporters suspected that this film was a disguised campaign document and that the persons depicted were all professional actors, a suspicion that was later confirmed.

The basic pattern of simple polarization in propaganda appears also in cartoons of a resplendent and heroic maiden representative of France (or the Allies or Western civilization) attacked by a terrifying gorilla wearing a Prussian helmet. One side must be depicted as wholly noble, the opposition as utterly depraved. St. George vs. the dragon, the forces of righteousness vs. the powers of evil, progress vs. catastrophe, humanity vs. barbarism—such is the elementary dualism of propaganda as contrasted with the complicated overlap revealed by a more scientific analysis. It is also elementary that the human need to structure any complex situation explains both the origin and the power of these oversimple stereotypes.

COMMON DEVICES

Clyde Miller, formerly a public relations counselor for certain utilities, and later Secretary of the short-lived but influential Institute for Propaganda Analysis (1937–41), has popularized the following 7 techniques of propaganda: (1) name-calling, (2) glittering generalities, (3) transfer, (4) testimonial, (5) plain folks, (6) card stacking and (7) band wagon. These devices depend for their effectiveness upon background conditions in the culture. A population that is not ready for the propagandist's stimulation will not be affected thereby. No amount of money, professional skill in publicity, or control of the channels of communication will suffice to make cannibals out of any modern industrial nation—the repugnance or taboo against this form of behavior is too deep-seated. Making Mohammedans out of contemporary protestant or Catholic Americans is all but impossible; the most that could be hoped for is a few converts and a greater tolerance toward Islam as a religious faith. By contrast, whipping up a war against Soviet Russia would be easy. Objective realities do rigorously confine the area within which the propagandist can work successfully.

It is comforting to know that the propagandist can accomplish only that which the cultural traditions and current status of the propagandees permit. He merely makes these traditions work for him by utilizing, magnifying, and circulating the terms and appeals already at hand. The common fear of propaganda is hardly an adaptive response to its presence since propaganda itself is usually little more than one group's insistent call to action—often dubious action, to be sure, but best combated, not by flight or simple assault, but by the strengthening of rational controls generally.

NAME-CALLING

Some of the formulas of name-calling plainly rest upon the layman's ignorance of semantics; they are influential because much of our language is full of multiple meanings and because many of our common concepts are vague and undeveloped. Pin an ugly label upon one's opponent and one has greatly crippled his influence with those who by prior training have learned to respond disapprovingly to this term. It is utterly false psychology but perhaps good botany or chemistry to claim with Shakespeare that a rose by any other name would be just as sweet. If a man is an Italian, that is ordinarily a fairly neutral designation, but call him a fascist (correctly or incorrectly, the consequences are similar) in an antifascist culture which equates that word with black reaction and his standing drops precipitately. A subtle blending of correct and incorrect uses of emotionally-saturated or defamatory adjectives and substantives is standard practice among all propagandists. The harm caused by libel and slander, occasionally leading to the nervous disintegration and even self-destruction of the victim, shows what a powerful instrument this can be.

GLITTERING GENERALITIES

The arousal of antagonism via name-calling has its complement in the evocation of approbation through glittering generalities. Pennsylvania New Dealers in 1936 called upon the citizens of that commonwealth to vote for "the man with a heart, the party with a soul." Such sloganizing need not be devoid of a factual basis any more than "making the world safe for democracy" was an utter fraud either in intent or temporary accomplishment. One unfortunate sophomoric consequence of becoming propaganda-wise may be the nihilistic assumption that all human purposes are rotten and that every cause conceals a latent or covert crookedness. Such exaggerated suspicion may lead to an inability to identify genuine values and to support them when necessary.

TRANSFER AND TESTIMONY

The transfer and testimonial technique are similar since both make use of prestige suggestion to support a particular position. Identification of the new specific purpose with a venerated symbol causes some of the manna to flow from the old to the new. It is like the pedagogical methods for making pupils interested in something—the new and neutral or even forbidding stuff is presented in connection with something that is already inherently interesting, the latter shedding its radiance over the other. When trial lawyers dramatically drag in God, motherhood, the Bible, home, and country to buttress their clients' case, they are relying on these dependable mechanisms to influence the jury in the desired direction.

Election placards and literature of all political parties make extensive use of the red, white, and blue symbolism. The national colors with their overtones of patriotic sentiment are the constant to which the special cause as a variable is ordered. Any person or organization possessing high status in a certain group may be effectively used in this way. The racial pride of the Negro in Joe Louis, the boxer, was capitalized by the managers of Wendell Willkie's presidential candidacy in 1940 when Louis lent his support to the Republican nominee. University presidents' names, especially if they are in "big league" institutions, are prized by pressure groups because of the popular deference these figures generally arouse and the untested assumption that such persons would support only worthy objectives.

PLAIN FOLKS

The plain folks device is a form of play-acting that helps to establish rapport between the propagandist (or the personal symbol of some cause) and the average man. Franklin Roosevelt was photographed eating frankfurters at a picnic and most candidates in campaign season find it desirable to be seen or reported fishing, pitching hay, riding a tractor, milking a cow, or a similar healthy or normal activity. These things are done, not because they are strictly relevant to the issue, but because they provide a favorable setting against which to project the message that is primary. The plain folks histrionics are preliminary to the more pointed propagandistic stimulus *per se,* serving to create confidence in and familiarity with the personalities who are delivering the suggestions.

CARD STACKING

Card stacking is another colloquialism referring to outright invention, suppression, or warping of data in the interests of some end. This may be done crudely or with such adroitness that it is difficult to detect the concrete act of chicanery involved. During World War I, a British army communiqué concocted the story that the Kaiser had referred to the English Army as "contemptible." Of course most of the people of Great Britain could have nothing but affection and admiration for their own kinsmen in the ranks even when they were beaten in battle. Hence, a boomerang effect was rightly counted on to intensify the popular hatred for the man who symbolized the enemy. Most Britishers today do not know that the Kaiser never said this, and they still pridefully refer to their troops as "The Old Contemptibles" (the fully documented account of this fraud is given in *Falsehood in War-Time* by Arthur Ponsonby, M. P., Chap. VIII, 1928).

Plain lies often work admirably—for a time at least. Newspapers unfriendly to unions may conceal the real grievances that actually provoked a strike and emphasize instead the rough behavior that may accompany it. The constant effort of communist or cryptocommunist officials to deny that certain front organizations are communist-controlled or dominated is a form of distortion that presumably assists in overcoming the name-calling handicap from which their activities would otherwise suffer by creating confusion among those not intimately informed about the facts.

BAND WAGON

The band wagon effect is a popular designation for the assumed tendency of people to want to be on the biggest or winning side regardless of the merits of the debate. Calling a wartime alliance the United Nations, even though it embraced rival allies and even though opposing Axis nations and certain neutrals were pointedly excluded, is an example of an effective appeal to majority-mindedness. The propagandist tries to create what the social psychologist has called an *impression of universality.* Few people are willing to be in a minority, unless that minority is favored in an important social way by winning prizes cherished in the culture. All other minorities are unpopular, and one seems to suffer in some way by being identified with them.

There is good experimental support (Marple, 1933) for the view that more persons change their opinions in the direction of what they believe is majority sentiment than in the direction of expert judgment. Evidently most persons become more secure in their beliefs when they find that these are shared by great numbers. I may be wrong but 100,000,000 fellow-citizens can't be! A sense of personal safety is enhanced by going over to the dominant side. However, Gallup, of the American Institute of Public Opinion, doubts whether there really is a true band-wagon effect in elections, since the division of opinion in national samples is frequently about the same at the end of a vigorous campaign as at the start, despite the wide publicity often given to straw-ballot percentages which are alleged to intimidate the minority.

The band-wagon influence seems to be of lesser magnitude than is commonly believed and apparently is operative only under limited social conditions. It may have little effect upon Republican and Democratic machines under our political duopoly, which rarely divide in anything more pronounced than a 60-40 ratio, but contribute seriously to handicap the growth of socialist, farmer-labor, and other third parties, which clearly must begin with smaller fractions.

PREJUDICE IN PROPAGANDA

The first thing which any skilled propagandist does is to equip himself with an intimate knowledge of the variety and intensity of the prejudices of the people he is seeking to influence. Next, the probable effect of each upon his special cause and its rivals is estimated. Then the message he aims to convey is tied in with each one of these concrete prejudices by various literary and artistic vehicles; strong emotional rather than critical intellectual appeals are preferred. The amount of demonstrable or probable enlightenment on the matter prevailing among the specific population subgroup being stimulated must be carefully considered. Finally, for maximum effectiveness the account that is dressed up must be widely distributed, particular care being taken to see that it reaches the bellwethers or molders of community opinion.

As an illustration of the plain appeal to prejudice, we may consider the March of Time film, *The Ramparts We Watch,* which was widely shown during the excited defense preparations of 1940–41. This had a staged scene depicting an American flag from a supposedly torpedoed vessel slowly sinking beneath the waves and another portraying an alleged German radio announcer declaring with contorted facial emphasis that "Americans don't make good soldiers." Such excessively strong stimuli reduced resistance at that time to conscription, vast armament expenditures, and active aid to Britain on a short-of-war basis, all effects which they apparently were intended to further.

Counter-propaganda depends upon the cultural level and other background conditions in the society concerned just as much as does propaganda. German counter-propaganda in the United States during the rising antagonistic tide fostered by official administration policies and other forces was planned to build good will toward things German by free distribution of beautiful colored photographs on a variety of cultural themes, reproduction of German medieval and Renaissance paintings of the Madonna and Child for circularization during the Christmas season, books and pamphlets on road-building, forestry, science laboratories, and other innocent or nonprovocative themes. Through fine printing and technically superior and interesting human photographs this German propaganda evidently indirectly tried to moderate the strong anti-Nazi sentiment prevailing in public and private quarters by suggesting that there were decent persons and institutions still operating within the Reich, rather than boldly attempting the impossible task of converting the already strong hostile attitude into a pro-German orientation. But the more favored British information services, from the point of view of American prejudice, concentrated on showing what a good government and culture the English have and how depraved their German opponents were; ostensibly they limited themselves to stating their case and then left it to their American readers (and listeners) to decide for themselves what form their assistance should take. Yet the suggestions implanted were always clear and strong and never ambiguous.

RESPONSE OF THE PUBLIC

REACTION TO PROPAGANDA

Most persons seem to recognize as propaganda that which they do not like for any reason and fail to identify as propaganda that which conforms with their apparently independent latent judgments about persons or situations. However, incidents do arise (as in defeated nations) where temporarily successful propaganda that has led to undesired consequences causes its victims to turn against and disparage those prominently identified with its manipulation. This is a delayed awareness of their own gullibility. In time of peace many citizens appreciate that they have been fed much gross propaganda during a past war period, but these same personalities may be victimized again in another war crisis because the social configuration usually changes sufficiently to permit an acceptable modification of techniques. The same holds for nonwar content unless revolutionary institutional transformations have meanwhile occurred.

Most of those who actively resist what is to them hostile propaganda come to accept without protest an objectively transformed situation that has become a *fait accompli* partly through such propaganda, apparently on the principle that "one can't unscramble scrambled eggs." The result is a constant race between the astute propagandist's ability to invent new and plausible appeals (largely because the old devices have lost their magic) and the relatively dull-witted and slow-moving public's ability to see through these mechanisms. A keener minority opposition may know full well what is happening and succeed in protecting the integrity of its own personality, and yet be impotent to change the course of events. Under the circumstances it is not hard to guess who will normally be ahead and who behind. This cycle of group acceptance-rejection or gullibility-disillusionment appears to be capable of indefinite repetition. Propaganda thus typically serves to perpetuate the social control of those who occupy the seats of power; less typically it may appear in the efforts of alert groups now out of authority to become its possessors. In other words, most propaganda is majority propaganda—an immensely significant fact. Propaganda is invariably a symptom and an instrument of the conflicts between ins and outs occurring in a disunited society.

MEASURING PROPAGANDA EFFECTS

The result desired by propaganda is an altered belief or action in a designated group. A certain fraction of the propagandist's effort must therefore be devoted to determining the kind and amount of change in public opinion he produces. It is at this point that the testing tools of the psychologist are conspicuously employed. Before any well designed propaganda campaign is begun, one must know the distribution of opinion on the issues with which one is concerned. Presumably, this opinion is unsatisfactory in some respect. For example, the William Allen White Committee to Defend America by Aiding the Allies, working to secure armed United States naval convoys for British ships in Atlantic waters (1941) might have discovered that on a certain date a sampling poll showed 30 per cent of the voters in favor of this quasi-belligerent step, 30 per cent opposed, and 40 per cent undecided. A further breakdown might have indicated the degrees of support and the intensity of opposition to the plan. If, in the judgment of competent persons, it seemed possible to change this unsafe distribution by certain appeals adequately publicized (like calling the move "sea patrols" rather than "sea convoys") a few weeks or months after the campaign a verifying or retesting poll could again have been administered. Assuming effective stimuli and the support of cooperating world events, the second (or final) poll might have indicated 50 per cent of the voters in favor, 25 per cent opposed, and 25 undecided. Whether this new distribution should be construed as a sufficient go ahead signal to public officials would depend upon the type of current citizen support they demand before venturing upon any significant move.

The uses and interpretation of such before-and-after data are usually very simple; certainly they are the safest and most exact guides now available. Above all, these procedures demonstrate conclusively that public opinion can be shaped by those in control of some or all of the media of communication and that, within the limits of the norms current in any culture, this important social phenomenon can be manufactured or made to order like so many machinery-produced articles. Ordinary commercial advertising unquestionably affects the results of market surveys made in terms of acquaintance with, or use of, specific products. Motion pictures, radio or platform speeches, cartoons and printed texts all create modification in thought and behavior. Just what this is qualitatively and quantitatively can now be ascertained with a remarkably high degree of precision by the opinionaire inquiries of contemporary journalism using suitably timed and stratified samples of mass response to reveal the changes resulting from the introduction of a large-scale stimulus.

VALUE TO SOCIETY

The utility of propagandistic activities for mankind as a whole is disputable. But there is no doubt that this is one of the ways in which the work of the world gets done. Much propaganda may not unfairly be described as quasi-criminal behavior. At any rate, there is no doubt that it is essentially psychological warfare—and therefore as good or evil as the institution of war itself. But current folkways do not look upon it as particularly pathological, perhaps because its vagueness, universality, and obscurity protect it from the resentment that is normally directed against other and more obvious antisocial conduct, such as forgery or petty thievery.

In perspectives of time and space, propaganda seems to be one manifestation of the competitive struggle for existence, perhaps doomed to disappear as cooperation among living creatures becomes the evolutionary norm. The ethical confusion of the current age complicates immeasurably the problem of estimating justly the services that propaganda performs. No one expects warring governments or bitter rival economic groups to deal scientifically or even judiciously with each others' antagonistic claims—a third party must usually be relied upon to approximate this; unfortunately, such a party is hard to find because almost everyone has some sort of prior attitudinal or emotional identification with the contending forces. Hence, noncontroversial themes are preferred by those who wish to pursue a less embarrassing type of research.

It could be argued that it is better for men to pursue certain group objectives, even if current community organization demands illusions and benign chicanery in so doing, than that they be left to drift along in a state of goallessness. Society needs directives and these the propagandist supplies. Truth, to be sure, is often sacrificed in the process, but then men live by other, and perhaps higher, values than this austere good alone seems to symbolize.

In a harmonious culture composed solely of well adjusted personalities, propaganda, as we know it, would probably have no place. Conceivably, the socially valuable part of contemporary propaganda could be incorporated in a beneficial system of informal education. Insofar as the aim of mature human effort is to approach such a state of affairs, it seems reasonable to expect that psychology, at least in its social, educational, and industrial branches, should eventually lead to the extinction of what we now know as propaganda and advertising instead of merely making them more efficient in terms of existing moral standards, as has been the case to date.

READING REFERENCES

Albig, William, *Public Opinion.* New York: McGraw-Hill Book Company, Inc., 1939.

Biddle, W. W., "Propaganda and Education." Teachers College, Columbia University, *Contributions to Education,* No. 531, 1932.

Doob, Leonard W., *Public Opinion and Propaganda.* New York: Henry Holt and Company, Inc., 1948.

Hartmann, G. W., "The Contradiction between the Feeling-Tone of Political Party Names and Public Response to Their Platforms." *Journal of Social Psychology,* 1936, **7**, 336–357.

———, "A Field Experiment on the Comparative Effectiveness of 'Emotional' and 'Rational' Political Leaflets in Determining Election Results." *Journal of Abnormal and Social Psychology,* 1936, **31**, 99-114.

———, "Immediate and Remote Goals as Political Motives." *Journal of Abnormal and Social Psychology,* 1938, **33**, 86–99.

———, and W. Watson, "The Rigidity of a Basic Attitudinal Frame." *Journal of Abnormal and Social Psychology,* 1939, **34**, 314–335.

Smith, Bruce L., H. D. Lasswell, and R. D. Casey (Eds.), *Propaganda, Communication, and Public Opinion, a Comprehensive Reference Guide.* Princeton: Princeton University Press, 1946.

5. PSYCHOLOGICAL WARFARE

JOSEPH W. EATON, PH.D.

Department of Sociology, Wayne University. Formerly: Personnel Research Section, AGO, Department of the Army; Psychological Warfare Division, Supreme Headquarters, AEF; Research Director of the Rural Settlement Institute.

DEFINITION

Men are needed to operate the deadly instruments of destruction, but the feelings and ideas of these men, their confidence in themselves and their cause, can turn the tide of battle. Experience in the fields of politics and government, business and advertising, and education and propaganda have shown that these human determinants of the mind can be subject to considerable purposeful control and manipulation.

World War I proved that it takes more than guns, ships, planes, and bombs to win a war. World War II saw many changes in military science, not the least of which was the development of methods of psychological warfare. Such methods were used on an unprecedented scale by all participants of the conflict.

Psychological warfare involves the controlled dissemination of ideas to help to achieve a military objective. It has a double function: first, the strengthening of the morale of the home front, and of Allied and neutral powers, and second, the destruction of the will to fight on the part of the enemy and potentially hostile neutrals. In times of peace the manipulators of the minds of men—politicians, advertisers, public relations experts, propagandists, religious leaders, salesmen, teachers, and others—are always confronted by a dilemma. They must either forego the use of certain techniques of persuasion which help them to obtain the immediate end in view, or violate prevailing moral codes. In psychological warfare, such conflicts seldom exist. The end justifies the means and all possible means are employed. Expediency rather than moral appropriateness is the decisive factor.

The role of ideas and of feelings in warfare has long been recognized. In every battle leaders make an effort to call out the fighting spirit in their men and frighten the enemy. Promises of fair treatment upon surrender have been used effectively ever since man had it in mind to rule others. The terrific increase in speed and facility of communication, made possible by such recent invention as the radio, the airplane, the telephone, the movie, and the mobile printing press, have given psychological warfare a new significance. It became possible to manipulate ideas, spread rumors, broadcast news, make suggestions, and arouse feelings among large groups of people, without any direct physical contact. It was therefore not surprising to find during World War II, that more than ever before, all combatants made extensive efforts to fight with ideas as well as with weapons.

The Germans took Austria and Czechoslovakia without a shot. Their propaganda lulled the rest of the world into a state of feeling secure or of fear, that made the Blitzkrieg of 1940 possible.

The Allied Powers also made considerable use of psychological methods. The Americans and British organized companies and a division of specialists among whom were writers, radio announcers, newspapermen, artists, printers, reporters, photographers, pamphlet-eers, censors, foreign language radio monitors, anthropologists, sociologists, and psychologists. These specialists, who in previous wars had little usefulness, functioned as a unit to turn ideas into weapons, to make thoughts do the work of swords, or at least to act as powerful supports of physically destructive weapons.

RELATION OF PSYCHOLOGICAL AND PHYSICAL WEAPONS

A comparison of psychological and other weapons of war is only possible on a superficial level. Small groups of patriots, poorly equipped with arms, but inspired by strong ideals of liberty, conquest, or revolution have often withstood and prevailed over large, well equipped armies which lacked an equally effective idea to give them morale. The American Revolution, the Irish battle for independence, the Russian Revolution, the partisans of Nazi-dominated

Europe, and the Zionist pioneers of Israel have demonstrated the combat effectiveness of ideas against superior force. And there are several inherent differences between psychological and physical weapons of war.

CULTURALLY CONDITIONED PSYCHOLOGICAL WEAPONS OF WAR

A captured German hand grenade is just as deadly when it is thrown into a crowd of German soldiers by an American infantrist as it is when used by the Germans. Ideas of psychological warfare, however, cannot be captured. They are a specific function of the culture of those who use them and those against whom they are used. In maintaining Army morale Americans made much use of "pin-up" girls in their newspapers for soldiers. However, when the Japanese dropped leaflets portraying pictures of female pulchritude over American lines to remind them of "what the civilians had at home while they were fighting in the jungle," the general reaction to this literature was, "These dirty Japs!" If anything, such leaflets made the American soldiers more fighting mad.

Another example of the cultural dimensions of psychological weapons was the promise of good and humane treatment of prisoners of war, which was used by all combatants during World War II. This idea, in the hands of the Americans and British, was a powerful propaganda tool against the Germans, but it was practically useless against the Japanese. The Russians and Germans used this form of appeal occasionally toward one another, but with relatively little effect. The idea of good sportsmanship and fairness in combat is in accord with basic Anglo-Saxon cultural traditions. Allied soldiers, by and large, could be depended upon to honor promises of good treatment, made by their propagandists to the enemy. The public at home could be depended upon to support such a policy of generosity toward those who laid down their arms. The idea of honorable surrender to superior force also has a long tradition among the Germans. In view of past contacts of Germans with the Western Allied nations over many years, such promises seemed reasonable, despite the efforts of the Nazi propaganda machine to change the German attitude toward them. With the Japanese, however, surrender appeals met with no response, for, from the point of view of Japanese culture, the idea of surrender is dishonorable. Fighting to the bitter end and suicide are far more desirable in terms of their cultural values than is the dishonorable experience of being captured alive by an enemy. Also, during the war there was less response on the part of American soldiers in making and honoring such appeals. To some Americans with strong feelings of racial prejudice, the Japanese were Orientals, members of a "yellow," or "colored" race, whom they considered inferior and inhuman.

Cultural determinants also are important to nations in the choice of general methods of propaganda to use in fighting one another. Germany, with her totalitarian culture made much use of terror, fear, and deception, techniques that could not be duplicated by the United States without seriously impairing the morale of many soldiers and civilians who had strong religious and democratic ideals. Germany and Russia could threaten and actually kill their hostages, for their culture gave support to the use of such means to achieve an end. Similar threats, which were made occasionally by the Western Allies, were less effective. Germans were less likely to believe that the threats would actually be carried out. Also British and American public opinion at best gave only hesitant and divided support to such measures.

CUMULATIVE PSYCHOLOGICAL WEAPONS

The range of psychological weapons is relatively independent of both time and space. Even an atomic bomb, with all its immeasurable power of destructiveness, affects physically only a few square miles during a specific period of time. No such limitations are attached to the combat-effectiveness of ideas. Through the radio and print, and by word of mouth, ideas can spread all over the world, over boundaries, through fortifications and even censorship. Their effectiveness, if the appeal is a strong one, can be sustained over long periods of time. An idea can hit both friend and foe; its effect cannot be isolated by either side. For example, despite vigorous German censorship (including the death penalty for spreading Allied propaganda) and ruthless German counterpropaganda about alleged Allied atrocities, the German Army and public became increasingly convinced that Americans were giving humane treatment to their prisoners of war. However, this idea could not be effectively hidden from outraged French, Belgians, and Russian peoples who often were hungry and almost frozen, while prisoners of war were relatively well fed and housed, peoples who had suffered grievously at the hands of the Germans and wanted revenge. Similarly, the ideals of the Atlantic Charter and the Four Freedoms could not be confined to the battlefields, but spread to the darkest corners of Allied colonial empires. While these ideals did much to undermine the morale of the enemy and strengthen the will to fight among all the Allied powers, they were equally destructive to many imperialist practices and privileges of the victorious power.

This cumulative effect can be an important limitation of psychological warfare. Ideas cannot well be used and discarded to suit temporary expediency. Their combat effectiveness depends greatly on other ideas which have preceded them and those that will follow. They must be plausible at the moment, and be sustained by future events. In view of this fact, psychological warfare planning was done by those on the highest level of Allied command echelons. It was controlled and coordinated directly by the Psychological Warfare Division (PWD) of Supreme Headquarters of the Allied Expeditionary Forces (SHAEF). In contrast, the Germans had a special Ministry of Propaganda, whose chief, Joseph Goebbels, had a direct voice in all important German policies.

The motto attributed to Lincoln, "You can fool some of the people all of the time, and all the people some of the time, but you can't fool all the people all the time," is perhaps the clearest summary that can be made of the cumulative characteristic of propaganda. Germany's failure to observe it had much to do with the ineffectiveness of many of her propaganda efforts during the latter years of World War II, just as scrupulous adherence to this principle brought dividends to the Allies. Regardless of the turn of military events, Germany's propaganda instruments were reporting a steady flow of glorious military feats. For a time, this deception maintained Germany's confidence in a possible victory. However, when Germany's forces, after retreating "victoriously" through Africa, Italy, Russia, and Poland, stood close to the German borders, this propaganda boomeranged. The Germans had to face not only the reality of defeat, but also the fact that they had been systematically misled. As a result they increasingly lost faith and confidence in their leaders as well as their system of government. On the other hand, the fact that the British Broadcasting Company and other Allied propaganda agencies usually reported the truth, built up a readiness among many Germans to believe all claims of the Allies. Also, the policy of the British government to admit frankly many reverses and difficulties had much to do with the high morale of the British people during the dark days of Dunkirk, and with their confidence in their government.

PSYCHOLOGY AND FORCE

Success of psychological weapons depends considerably upon the actual or potential military and economic forces that are available. A propagandist can threaten, malign, deceive, and make appeals; sometimes he can win battles by bluffing without firing a shot, as did the Germans in Denmark. However, without actual or potential power, propaganda is a hazardous enterprise. In any long conflict, bluffs are likely to be called; claims must be borne out by future events, or they are likely to have the effect of a boomerang. German threats in 1940 to bomb Amsterdam or Paris were as effective as an actual bombardment would have been. They brought about the surrender of the cities because the French and Dutch authorities were weak, and they believed that these German threats could be carried out. Similar threats to blockade England and America, which could not be backed up effectively, served only to expose the relative weakness of the German navy.

It is practically impossible to separate the relative effectiveness of military, economic, and psychological weapons which are combined in total war. Not only is the use of techniques of psychological warfare a function of the military and economic strength of the combatants; military and economic weapons also have psychological effects. The first atomic bomb dropped on Japan did not merely wipe out most of the city of Hiroshima; it greatly reduced Japan's will to continue fighting—much more so than would perhaps have been warranted by the actual physical damages of this new weapon. Similarly, American relief supplies to Vichy France did something beyond filling the bellies of starving children. It gave many downhearted Frenchmen hope in America's strength and good will.

TYPES OF PSYCHOLOGICAL WARFARE

The goal of psychological weapons is the same as that of other military weapons: to protect friendly forces and to bring about the defeat of the enemy. During World War II, the Allied forces conducted their propaganda on two levels: the strategic and the tactical.

STRATEGIC PROPAGANDA

In general, strategic propaganda was intended for large groups of people, disseminated over wide territories, including wide military fronts, the enemy's rear areas, and neutral countries. It had long-range objectives, requiring careful planning and coordination. It involved basic political decisions. This type of propaganda was difficult. It had to appeal to as many and antagonize as few interest-groups as possible, of which the potential audience was composed. However, the "loss of customers" could never be avoided completely. For example, the Allied agreement to make peace with the Axis only on the basis of unconditional surrender induced many Germans and Japanese, who were tired of war, to fight a little longer. This negative effect was perhaps offset by the fact that it prevented any possible conflict among the Allies concerning the terms under which an armistice might be granted; put an end to many enemy speculations that a prolongation of the war would elicit favorable terms in a negotiated peace; and increased the chances for a permanent world peace by making the complete abolition of fascist regimes a condition for ending hostilities. The strategic propaganda of Germany also scored certain outstanding successes in the early phases of the war. Scandinavia and the Lowland countries, lulled into a false sense of security through neutrality, held back the potential support that they might have given the Allies until the Germans were ready to conquer them with vastly superior force.

TACTICAL PROPAGANDA

Tactical propaganda was intended for small groups of people, disseminated over specified areas of a military front, the enemy's home front or neutral countries. It had specific and immedate objectives, such as the surrender of a certain area or the sabotage of a certain installation by pro-Allied sympathizers behind German lines. This type of propaganda required detailed intelligence of the specific situation and all its possibilities. No promises or statements were made which were in contradiction to the general strategic propaganda approach. Tactical propaganda, because it usually appealed to a single audience and interest group, often produced remarkable immediate results.

WHITE AND BLACK PROPAGANDA

Tactical propaganda was carried on in two separate branches—*white* and *black* propaganda. In white propaganda the origin was clearly stated; in black propaganda, the origin is not merely hidden but falsified. White propaganda included such activities as radio programs, leaflets, newspapers and official news-agencies (such as the British Broadcasting Company and the Tass News Agency) which clearly spoke as the responsible agent and voice of the government they represented. Black propaganda was intended to confuse the enemy, by posing as an enemy radio station or leaflet starting false rumors, or printing statements purporting to come from enemy sources. The Russians, for example, printed a complete imitation of a Nazi songbook and distributed it through their agents among German soldiers. In the middle of the book were a few soul-stirring anti-Nazi songs. The British and Americans kept two German radio stations on the air after the invasion of Europe, which claimed to be operated by German and Nazi soldiers within German-held territory. These stations did much to undermine German morale, were very popular among German listeners, and were accepted by many of them as bona fide German stations.

TECHNIQUES

INSTRUMENTS

Psychological weapons were communicated through written and spoken language, as well as symbols and pictures. In strategic propaganda, radio stations played the most important role. There also was considerable use of printed leaflets, newspapers, and pamphlets, dropped by planes over enemy lines, to convey messages through texts, photographs, maps, and cartoons. Motion pictures were also used in propagandizing friendly, neutral, or conquered enemy populations. In tactical propaganda, loudspeakers and portable short-range radio stations were important channels of communication. Mobile printing presses, capable of printing millions of leaflets or posters, quickly furnished propaganda material needed to exploit an urgent or favorable tactical situation. The British and American Armies had special Mobile Radio Broadcasting Companies, fully equipped to conduct every type of tactical propaganda in the front-line areas. They were staffed with prisoner of war interrogators, research men, foreign language radio monitors, and announcers, script writers, newspapermen, radio technicians, and printers. Leaflets were dropped by plane or shot by artillery guns in special shells over enemy lines. The special technical equipment and the professional training of these propaganda companies permitted speedy operation which was necessary to create or prevent panics, make surrender appeals, give instructions and disseminate up-to-date news and information.

METHODS

Psychological warfare techniques made full use of all available knowledge of the nature of the social target. Anthropologists, psychologists, sociologists, and historians reviewed the folkways, mores, emotions, and aspirations of each group to furnish general guidance to the propagandists. They prevented attacks on the Japanese "God" emperor or the use of the term "coward" in connection with German troops, since attacks on these symbols would have strengthened rather than weakened enemy morale.

On the basis of careful studies of the conditions and moods of the audience, the following methods of approach were most commonly used:

INTENSIFICATION OF EXISTING FEELINGS. Propagandists realized that it is difficult to convert people from one point of view to another, particularly in a war, when everything is defined as either "for" or "against." Emphasis was therefore put on intensifying existing sentiments, that would help one's cause. Thus the Germans used anti-Semitism in the belief that this prejudice was already widespread among many groups. Slogans, such as "Britain will fight to the last Frenchman," were based on traditional suspicions between the two nations. For the same reason, American propagandists did not attempt during the war to convert Germans into philo-Semites or lovers of democracy. Prominent German exiles, who were unpopular in Germany, were not used to make appeals to their fellow-countrymen. Even Hitler was spared from attack as long as he was found to have the confidence of too many Germans. Instead, unpopular Nazi leaders and policies were emphasized. Factual accounts of the ruinous cost of the war to Germany, corruption among certain Nazi leaders, and the hopelessness of Germany's struggle against the power of the rest of the world were stressed, in the knowledge that these sentiments were already present among many Germans.

SHOW OF KNOWLEDGE. This method was based on the realization that a show of knowledge is always impressive, particularly when it involves facts that are considered important, secret, and essential. Troops on both sides were often greeted by radio messages or leaflets from the enemy, informing them that their exact strength and position was well known and that the enemy was well prepared for their arrival. Such messages often had negative effects on the morale of the troops, particularly if they knew themselves to be in a weak position and then found that the enemy knew it too. It is here that the findings of espionage were often put to effective use. Show of knowledge of intimate minor details induced many prisoners of war to tell all they knew, for fear that their interrogators were so well acquainted with the situation that efforts to mislead or deceive would be futile, if not dangerous. The realization that there were spies among them, served to decrease their faith in their comrades and to increase their respect for the power of the opponent.

TRUTH. The Allies perhaps made "truth" their most potent tool of psychological warfare during World War II. Propaganda agents had strict instruction never to resort to lies, except with special permission from top-level commands. It was realized that lies, when discovered, tend to discredit all future pronouncements. However, the decision to use nothing but actual facts in propaganda, did not mean impartiality. By selecting factual data, by omitting or "playing down" unfavorable situations, every attempt was made to influence the enemy in the desired direction. Thus Allied news reports would give prominence to actual Allied superiorities, expose fictional German "secret weapons," and refrain from making promises or prophecies that could not be kept. They ignored religious or political conflicts among the United Nations, except those that were already too well known among the Germans to be ignored and they reported the favorable settlement of these differences.

DECEPTION. German propaganda made much use of invented news reports, slander, rumor, and outright falsification. This method gained them many short-run advantages, but in the long run contributed to the loss of confidence of their audiences. Foreign workers were lured to Germany by false promises of good treatment. The resentment arising from this deception had much to do with turning the millions of foreign workers in Germany into an active instrument of sabotage. Many German soldiers deserted when they discovered that their leaders had constantly misrepresented the military situation, although for a time these deceptions kept soldiers fighting despite the military hopelessness of their position. For these reasons, Allied propagandists used deception very sparingly, except in their black propaganda. Since black propaganda was disguised to originate from German sources, the discovery of its occasional falsehood would discredit German rather than Allied prestige.

FEAR. Threats and terror were used widely as methods of psychological warfare. Germans were kept fighting by their leaders through threats of the horrible fate that would await them and their families under Allied control. Also, because all attempts at desertion were punished by shooting, toward the end of the war many Germans continued to fight largely because they believed to do so was safer than giving up arms. In other words, the Gestapo was closer than the Allies. Fear also held millions of Europeans under German rule, people who were ready to rebel at any moment. These fears often were strengthened by acts of terror, such as the shooting of hostages, extermination of entire villages and families, concentration camps, and other inventions of torture. The Allies also made considerable use of fear. Many a German city surrendered in 1945 when threatened with a devastating bombardment by Allied guns and planes.

PROMISES OF REWARD. The cooperation of people can often be secured most easily by promising them goals toward which they are striving. Such promises can be made with relative ease to persons fighting on the side of the propagandist. Thus the Germans held out to their people a rosy future as the powerful rulers of the world. American soldiers knew that valiant fighting would bring them earlier peace, a trip home, promotions, praise, medals, and other rewards. One of the great shortcomings of Germany's propaganda in Europe was its failure to offer sufficient inducements to the large masses of the conquered nations to cooperate in the building of Hitler's New Order. The Allies, on the other hand, were careful to offer actual rewards to their enemies in return for surrender. They promised fair treatment under the rules of international law, peace, liberation from dictatorship, and opportunity to participate at a later date in the operations of the United Nations.

POSTWAR PSYCHOLOGICAL WARFARE

When an armistice is concluded, soldiers lay down their arms, but not the propagandists. However, there usually is a change in their ultimate objective. During a war, the objective is to defeat the enemy at any price and with any method. After the cessation of hostilities, the objective shifts to enhancing the prestige of each power. The fear of possible future armed conflicts makes it seem advisable to make as many friends and influence as many people as possible. Less emphasis is placed on methods involving fear and more stress on mutuality of interest. Perhaps one of the most disturbing signs to those who wish to preserve peace is the fact that several of the great powers, who fought this war, continue to maintain government-manipulated propaganda agencies, operating to a considerable extent under the same rules that they used during the war. This statement applies not only to all totalitarian governments, but to many nontotalitarian states as well. The daily newspapers are full of statements by government officials, charges and countercharges, rumors, slanders, and fabrications which give evidence to the fact that there is still too little confidence among the nations of the world to communicate with one another on the basis of mutual trust. The armistice in psychological warfare is still to be declared.

The importance of international understanding and confidence is well expressed in the preamble of the United Nations Economic and Social Council (UNESCO), which was established to promote peace through these methods: "Since wars begin in the minds of men, it is in the minds of men that the defenses of peace must be constructed."

READING REFERENCES

Doob, Leonard W., *Propaganda; Its Psychology and Technique*. New York: Henry Holt and Company, Inc., 1935.

Farago, L., and L. F. Gittler (eds.), *German Psychological Warfare; Survey and Bibliography*. New York: Committee for National Morale, 1941.

Hitler, Adolf, *Mein Kampf*. New York: Reynal and Hitchcock, 1940.

Sington, Derrick, and Arthur Weidenfeld, *The Goebbels Experiment*. New Haven: Yale University Press, 1943.

6. UNEMPLOYMENT AND PSYCHOLOGY

M. R. TRABUE, PH.D.

Dean, School of Education, Pennsylvania State College. Formerly: Director, Occupational Research Program, US Employment Service, Bureau of Educational Service, Columbia University, and Bureau of Educational Research, University of North Carolina; President, American Educational Research Association, National Vocational Guidance Association, and Council of Guidance and Personnel Associations.

No one would claim that all the problems of unemployment can be solved merely by applications of psychology. No informed person would deny, however, that any effective solutions for these problems must be in close harmony with the psychology of individual and group living. Psychological services are not a panacea for unemployment, but they have an extremely important contribution to make in the alleviation of this modern social ailment.

The most important contribution of psychology for the attack on problems of unemployment is its emphasis on individual differences. Unemployment in a community, state, or nation involves many unemployed individuals who differ in certain important respects. They differ, for example, in the relative amounts of activity and rest they require or desire. The people of one community may be perfectly happy and contented with a percentage of unemployment that would disturb profoundly the citizens of another community. The successful workers in any occupation possess in common a certain combination of attitudes, interests, aptitudes, and abilities which differs in certain respects from the combination that characterizes successful workers in other occupations. Determining the nature and extent of such differences among individuals, among occupations, and among communities, and providing for such differences, is the major task of the psychologist who is faced with the problems of unemployment.

INDIVIDUAL UNEMPLOYMENT

THE UNEMPLOYED

The popular conception of being employed involves working regularly each day for wages previously agreed upon. Thus, anyone who does not on a given date have a steady job that brings in a regular income would be unemployed. An artist, a housewife, or a doctor might be working long hours every day, but each might still be classified as unemployed so far as having a steady income from an employer is concerned.

No definition of employment separates clearly those who are from those who are not in need of psychological help in the solution of unemployment problems. Some persons feel little or no need of an assured cash income in order to consider themselves fully and profitably employed, while others who have regular incomes may consider themselves the unfortunate victims of unemployment. Whether an individual is or is not suffering from unemployment depends to a considerable extent upon his own standards and ideals. This fact emphasizes the importance of psychological services in alleviating the distress caused to the individual by unemployment.

PSYCHOLOGICAL SERVICES

Because unemployment means different things to different people, the psychological service needed by one unemployed individual may be quite inappropriate for another. No prescription can be adequate for all who suffer mental distress from unemployment. Each person must be fully understood before appropriate guidance activities can be chosen and administered. This requires of those who would assist the unemployed a highly specialized training in examining and analyzing individuals, as well as accurate knowledge of current industrial and labor conditions and of the alleviating agencies and resources that are available.

EXAMINATION

One of the first questions to which a psychologist must secure an answer if he wishes to help an unemployed person is, Why is this person now out of a job? The unemployed individual himself is almost certain to have a ready answer, but it may be far from the correct one. "My boss didn't like me," or, "I quit to look for a better job," may be the man's own explanation. The employer's report—"The man is lazy" or "He is no good"—cannot be accepted at face value. The psychologist must secure a far more discriminating answer to this question based upon accurate information gained through interviews and examinations.

He studies the man's occupational qualifications, emotional characteristics, social relationships, vocational ambitions, physical condition, and character traits, particularly as these are related to the jobs at which the man has worked or would like to work. He uses every means available for the acquisition of accurate information. Interviews are held with those who can give reliable information about the unemployed man. Records regarding the man and his family are collected from schools, stores, industrial plants, and social agencies, Tests and inventories of knowledge, skills, attitudes, interests, and aptitudes are administered, and the results collated with the data obtained by other means. Through careful study of these evidences, the pattern of the unemployed individual's occupationally important characteristics is identified.

It should then be possible to compare the results of these studies with the combinations of traits that are known to contribute to success in the occupations which the unemployed individual has tried or would like to enter, and to determine thereby the most probable reasons for his unemployment. It is well, in some instances, to discuss this analysis frankly with the individual, so that he may gain a clearer understanding of the factors behind his unemployment and be prepared to avoid repeating the same errors. In other instances it is better to give the man no hint of the conclusions reached, but to offer specific suggestions and help that can be used at once without much reflection or self-analysis. Guidance procedures must be intelligently adapted to the individual's character, personality, intelligence, and needs.

UNEMPLOYED ATTITUDES

The attitudes of an unemployed person must be made the subject of careful study by the psychologist. Most unemployed persons are quite certain, especially in times of widespread unemployment, that the one thing they need most is a job—any job that will produce a steady income. This is particularly true when the individual has been out of work for a considerable length of time. The financial return may actually be much less important, however, than the social prestige of "having a job" or the physical release that comes from activity itself. Merely having a place at which one must report and remain for a certain number of hours each day may give real satisfaction and comfort.

There are probably some individuals whose pride and character suffer little when they "go on relief" or are assigned to "made work." Others suffer great mental distress under such circumstances. Those who are sensitive about it should be given appropriate psychological therapy. The mental and emotional strain caused by losing a job may in itself be sufficient to modify the individual's whole personality. A normally cheerful person may become extremely gloomy and secretive. Efforts to relieve the distress of an individual's unemployment should always be made in terms of his own distinctive interests, attitudes, character, and needs, and should be concerned quite as much with his mental health as with his empty stomach.

This distress of mind arising from loss of a job, or from inability to find one, often develops through fear of what others will say or think when they learn of it. There was plenty of work to be done during the early history of the nation, even though it may not always have been highly remunerative. Laziness and idleness became, under such circumstances, almost synonymous terms. An able-bodied person who was not doing his share of the community's work was not considered a good citizen. Because these social attitudes still persist, many unemployed people try to keep their friends from knowing that they have lost their positions. A related reaction often leads the unemployed to refuse to apply for public relief until their economic situation has become desperate. Many keep trying to find work for at least a year after being "laid off." Persons who have given up trying and have accepted public relief are sometimes so uneasy about their own social status that they try to persuade other unemployed persons to stop trying and to join them on the relief rolls.

GENERAL UNEMPLOYMENT

Alleviation of unemployment as a social ill requires a realistic understanding of the changing social, economic, industrial, and occupational conditions affecting individuals. Psychologists may contribute substantially to a general understanding of the psychological implications of these changes through lectures, discussions, or printed articles. It is useless to advocate the discontinuance of invention and mechanization in industry, or to tell unemployed people that their difficulties are the result of substituting machines and conveyor-belts for human labor. As long as manufacturers compete for business, each is inclined to try to make his product better and to produce it at a lower cost. Mass production methods and automatic machines reduce production costs and make it possible for more people to enjoy a greater number of conveniences. No man-made law can repeal the desires of human beings to possess a larger number of these modern devices and comforts in a world that is flooded by extravagant and insidious advertising.

UNEMPLOYMENT CONDITIONS

A man who has been thrown out of employment because his employer invented or purchased a machine to do work that was formerly done by hand is not going to be greatly comforted by the assurance that in the long run more people will be employed than before. It may be true that several times as many persons are now engaged in the garage business as were formerly employed in the livery-stable business, but many a man who was successful in caring for the needs of horses found himself totally unable to learn to repair automobiles. Invention and the routinization of productive industry do lower the prices of commodities and thereby create new demands, which may result in the ultimate employment of more persons than ever, but the individuals thrown out of work are not always the ones employed later to operate the machines.

The introduction of more automatic machinery is certain to continue to make jobs more repetitive and routine, less complicated, more readily learned and less interesting to the worker who possesses varied skills. The hours of labor will continue to decrease, except in times of special emergency. With this shortening of the time spent at work and this simplification of the tasks performed, each worker will find it more and more necessary to discover opportunities for the exercise during his leisure hours of special skills and knowledge no longer required in his regular work. Psychologists may be very helpful in aiding workers to select the leisure-time activities in which they can secure satisfying exercise for these special skills and interests not required in their regular employment.

SOCIAL CONTROL

Seasonal lay-offs and general economic depressions must be treated as products of man's sluggishness in creative thinking and cooperation in social and economic programs. Greed, selfishness, and lack of willingness to sacrifice present personal gain for future general welfare are certain to produce social conflicts, including war, waste of resources, and general unemployment. There are employers who plan far enough ahead to be able to build up their stocks of finished goods during slack seasons and thus keep their employees busy throughout the year; but many employers prefer to rush production during periods of strong demand for their goods and thus to require their employees to take the risks of overproduction and economic depressions. Government regulations may ultimately correct this situation to some extent.

Labor unions have been an effective means by which workers have been able to force heedless employers to consider their needs. The power of the unions has sometimes been used for selfish ends by their officers, with the result that the workers have often found themselves unemployed for unnecessarily long periods. Selfishness is quite as likely to dominate the activities of a worker as of an employer. Both the worker and the employer must learn that neither can long be exploited without doing injury to both.

EDUCATION OF SOCIAL ATTITUDES

In the relatively simple social and economic conditions of pioneer days, the employer and employee were well acquainted personally and each felt a definite responsibility for the welfare of the other. As industries increased in size and complexity, however, the personal relationships and feelings of mutual responsibility grew weaker and more difficult to maintain.

Stockholders insist today on receiving the largest possible returns on their investments. Boards of directors require their executives to produce maximum profits. The managers try, therefore, to get the work done with the lowest possible expenditure of time and money. The relationships between employers and employees have become so distant, so impersonal, and so difficult to understand that only through definite instruction are they likely to be recognized at all.

Although there are as yet relatively few teachers in our public schools who recognize and attempt to meet their instructional responsibilities for clarifying the mutual dependence of each group upon the other, it is only through public education that we may hope to develop and maintain any clear understanding of such complex problems as those of employer-employee relations.

The schools are maintained by the state primarily for the purpose of improving the effectiveness of the social, civic, and economic programs that affect the common welfare. The primary task of public education, therefore, is to develop in all citizens those understandings, attitudes, and skills which will promote more effective cooperation in solving their common problems of living and working together.

Whether an individual will use his knowledge of arithmetic, his skill in writing, or his training in chemistry to promote the common welfare or to seek a selfish advantage over his fellow citizens depends chiefly upon his attitudes toward other people. The most fundamental obligation of the public schools, therefore, is to develop attitudes of cooperation and mutual respect for other persons and groups.

It would be quite unrealistic to believe that even the most effective possible procedures in education could completely eliminate self-interest in citizens. The complexity of modern life and the terrible damage which an irresponsible group can now inflict upon entire nations makes it necessary, however, that all be taught to be socially responsible and cooperative. Unless all groups do recognize their interdependence, develop attitudes of social responsibility and habits of cooperation in the solution of common problems, and organize to prevent persons who are irresponsible from being allowed to exploit others, our entire civilization may soon be destroyed. Public education is the chief hope we have for survival in the modern world, and psychology is the science which must supply the procedures by which the desired attitudes, understandings, and habits can be most effectively developed.

OCCUPATIONAL ADJUSTMENT

The vocational psychologist seeks to discover the particular combination of physical, mental, and emotional characteristics possessed by those who enjoy their tasks and succeed best at each different type of work. He also seeks to discover the combination of characteristics possessed by each candidate for work, and then to help the candidate to prepare for and to obtain work of a type in which he will be as typical as possible of the best workers. The techniques involved here have been developed chiefly through the activities of vocational psychologists in schools and industries.

To provide a worker with maximum satisfaction, a job should not only bring in a reasonable financial reward, but it should exercise as many as possible of his special skills and interests, and it must not require the exercise of abilities that he does not have and cannot readily learn. The progressive substitution of routine machine-tending jobs for the older-type jobs that required resourcefulness and a wide range of special abilities will cause more and more workers to be dissatisfied with the monotony of their work. This result can and should be partly offset by seeing that the more versatile workers are given appropriate additional training, with adequate opportunities to transfer to more complex and more satisfying jobs. Each change in the process of production should be accompanied by careful readjustments of workers and jobs throughout the plant. The techniques of the vocational psychologist are important tools whenever such studies and readjustments are to be made.

It is evident, in view of the constantly changing character of jobs, that a youth can no longer choose a trade with full confidence that he will be able to follow it for a lifetime or even for a single decade. The public schools should supply each youth with a reasonably clear understanding of his own abilities, aptitudes, and occupational possibilities. A youth should be taught to expect to change occupations a number of times during his lifetime, but always to hold, as closely as possible, to the type of work that is appropriate to his own particular combination of occupationally important characteristics.

AIDS TO SELF-GUIDANCE

The work one has already done and enjoyed doing is an extremely helpful clue to the type of work to be sought. An exploratory industrial arts course in school should develop sound attitudes toward work and basic elements of skill. It should help the psychologist to identify the type of work or family of occupations in which the largest amount of satisfaction and success may be expected by the individual.

Real progress has been made by psychologists in identifying "job families" and in showing how different jobs are related to each other in the sense that they require the same or similar combinations of worker characteristics. A youth's vocational plans should be developed in terms of his own capacities and interests, but they should contemplate his remaining in an appropriate family of occupations rather than in a particular job, trade, or industry. Each change of employment may then require the learning of a number of specific applications of his qualifications to the new plant or industry, but it will not involve the acquisition of an entirely new set of skills and interests.

THE CONTRIBUTION

Psychologists have developed useful instruments and techniques for evaluating important characteristics possessed by individual workers—both those who are successfully employed and those who are unemployed. Experience has shown that these psychological tools may be extremely helpful in aiding unemployed individuals to make desirable occupational adjustments as well as in aiding employers to select qualified workers.

Mass unemployment is a phenomenon requiring cooperative social and economic thinking, as well as a thorough understanding of human psychology. Its alleviation depends upon adequate understanding of these matters by both employers and workers. No effective planning of social programs in this field can ignore differences in the qualifications of available workers and differences in the requirements of jobs. It is at this point particularly that psychology becomes an essential tool in the continuous process of individual adjustment and social reconstruction.

READING REFERENCES

Bingham, W. V., *Aptitudes and Aptitude Testing*. New York: Harper & Brothers, 1937.

McGregor, D., "The Attitudes of Workers toward Layoff Policy." *Journal of Abnormal and Social Psychology*, 1939, **34**, 179–199.

Paterson, D. G., John G. Darley, and R. M. Elliott, *Men, Women and Jobs*. Minneapolis: University of Minnesota Press, 1936.

Roethlisberger, F. J., and W. J. Dickson, *Management and the Worker*. Cambridge, Mass.: Harvard University Press, 1939.

Shartle, C. L., *Occupational Information*. New York: Prentice-Hall, Inc., 1946.

Watson, W. F., "Strikes and the Human Factor." *Human Factor*, 1935, **9**, 150–157.

Wyatt, S., "Workers and Machines." *Occupational Psychology*, 1939, **13**, 249–257.

Yoder, Dale, *Personnel Management and Industrial Relations*. New York: Prentice-Hall, Inc., 1942.

7. PSEUDOLOGY

JOSEPH JASTROW, PH.D.

Deceased. Formerly: Professor of Psychology, University of Wisconsin; First Secretary (1892) and President, the American Psychological Association, Inc.; President, the Eastern Psychological Association; pioneer in popular writing on scientific psychology and its application.

All knowledge is subject to error and distortion, and characteristically to prejudices and fallacies which are congenial to primitive and undisciplined thinking. Pseudology serves as a term for the array of false beliefs that thus arise in all cultures. The pseudologies of any science are of direct interest to psychology because they represent untrained modes of thought and belief. Thinking is natural, but right thinking is an art, acquired slowly and with difficulty.

Many pseudologies invite quackery. Quackery, which is the practice of pseudologies, arises from exploiting the ignorance of the masses and playing upon their desire to learn of their powers, fortunes, and fates. But likewise quackeries arise as counterfeits of true knowledge. They have been stimulated by the rapid growth of psychology, especially in the fields of personality analysis and vocational guidance. Psychology, bearing so closely on mental health, shares with medicine in the tendency to offer quick cures and panaceas.

Pseudologies arise in response to human needs. Prominent is the craving for the dramatic that bends truth to wishful imagination. The age of fable preceded that of science, and observation finds a rival today in dramatic imagination. The irregular stream of pseudologies arising throughout the ages derived from two sources: (1) the uncritical folk-mind devising homespun beliefs in part summarized as superstition; and (2) the learned and doctrinal mind that reaches conclusions on such evidence as is available and in favor in each age as a part of the effort to correct the errors of its predecessors. Modern pseudologies flourished particularly in the eighteenth and nineteenth centuries—some new and others revivals of ancient error. The doctrinal pseudologies reflect both folkways and the attempt to adapt belief—often by crude counterfeits—to the patterns of accredited science.

ANIMISM

The ways of thought characteristic of the untutored, the natural, and the primitive mind persist; they have a vast, if not profound, influence on human progress. When these ways of thought are elevated to doctrinal levels they become pseudologies. Sympathy is so strongly felt in oneself that primitive man projects it into the order of things generally. The simple folkminded find it congenial to plant corn under an increasing moon so that it shall grow with the lunar growth, and to cut hair and nails as the moon is decreasing, and find it poetic and natural that Barkis "goes out with the tide."

FABULOUS ZOOLOGY

We have as a result a zoology of fable; of griffins, centaurs, unicorns, and basilisks; of the phoenix that arises from the ashes and the sea-serpent that swallows boat and crews; and of the endowment of pet animals with high intelligence, such as talking dogs and calculating horses.

EARLY PSEUDOLOGIES

Early pseudologies elaborated the doctrine of sympathy and such related notions as attraction and repulsion which they found in intimate human relations. They saw the same humanized forces operating in cosmic movements, in the seasons, and in animal and mineral nature. The doctrinal mind sought and found in nature the signs or signatures indicating service for human welfare and extended it by use of symbolism and correspondences. There was thus created a way of animistic thinking which survives.

ILLUSTRATIONS

Names of plants reflect such origin, as eyebright, liverwort, heartsease, quaking grass for ague. Moreover, pine cones are for toothache since they resemble the teeth, and a diet of walnuts is prescribed for brain trouble because the convoluted kernel resembles the brain and the hard shell the skull. Such inventions as the Powder of Sympathy and the Weapon Salve of Sir Kenelm Digby (1640) are illustrations. As the salve was rubbed on the weapon, the wound which had been inflicted by it healed. It led to the saying, "To take a hair of the dog that bit you"; to casting spells by hexing on a lock of hair or some belonging of the victim; or pricking a lemon with thorns—a substitute for the enemy—and thus inflicting tortures on the person of one's enemy. It led to the belief that as a buried glove of the enemy rotted so would the body of the victim.

As science advanced, the pseudological systems assumed semblances—spurious counterfeits—of scientific evidence. The persistent current that runs through them all from ancient to modern, from the jungle to the metropolis, is the intrusion of the animistic or subjective into the scientific or objective, the belief that somehow the world responds to human interests, and that wisdom lies at the end of the rainbow of wishful thinking.

PARENT PSEUDOLOGIES

ASTROLOGY

Typical of the pseudological complex is the patriarch of all pseudologies, astrology. It is a fanciful structure which includes (1) the fact that man gave the names of gods to the major planets—Jupiter, Mars, Venus, Saturn, Mercury—and ascribed to them human qualities; and (2) the creation dividing the calendar year into twelve "houses of Heaven," named after prominent constellations and designated by animal and anatomical "resemblances." Also, on the basic assumption that the hour of birth is decisive for human fate, it reads the qualities and fortunes of human lives by the alleged influences of the celestial bodies. No more illogical and superstitious system of notions ever could have originated from a learned but confused brain than are epitomized in the horoscope. Yet this pseudology has developed into a lucrative quackery, perhaps just because it does take such a flight from reality.

Astrology was not without progeny. The most prevalent was cheiromancy and its modern revival called palmistry. Palmistry operated by transforming a still more elaborate order of human fates to the prominent anatomical creases of the palm. Other pseudologies arose of the same ingenious invention. All moved into quackery as signs, meanings, interpretations, and readings followed, with fanciful assignments appealing to pseudolearned minds completely ignorant of the logic of science.

Astrology violates every conception of the nature of the universe and all scientific procedures to determine fact. Two psychological conclusions are illustrated by it: (1) wishful thinking gives way slowly to disciplined habits of thought and (2) the logic of verification of fact is a weak spot or blindspot in popular thinking today.

ALCHEMY

Equally ancient and equally characteristic is the other great parent pseudology, alchemy. Its appeal is to material wealth—gold. There is no psychical factor in alchemy, but in the use of magic expressed as transcendence over physical powers, it reflects the typical family features of all ancient and medieval pseudologies.

As the predecessor and rival of scientific chemistry, alchemy became the prototype of the magical, the secret, and the occult arts. The alchemist was a long-range successor of the medicine man and the shaman, sharing with him the province of wonder-working. When he exercised his arts for worthy purposes, it was white magic; when by the aid of the devil he wrought evil, it was black magic. In the course of his practice, one specialty became of great importance—the transformation of the baser metals into gold. Inevitably, in pretending to do so and in furnishing an exhibit in proof, ingenious quackery was practiced.

The elaborate procedures of alchemy were based upon abstruse speculations concerning the processes of conversion and sublimation, attraction and repulsion, and other mixed physical and psychological reactions, e.g., correspondences and sympathies. The ultimate search was a wish-goal, called the "philosopher's stone," which would confer all boons and solve all mysteries: health, wealth, perpetual youth, the elixir of life, universal knowledge, and the power to predict and control fate—in brief, the rending of the veil that stands between man's limitations and the secrets of the universe.

The appeal of alchemy was ever to the doctrinal mind because amateurs could not dabble in it for lack of sufficient knowledge. Because its pursuits are so closely related to legitimate physical chemical problems does not establish it as scientific. Alchemy is based upon premises and procedures now classed as error of pseudology and pretentious counterfeits of accredited science.

PSYCHIC TRANSCENDENCE

A major desire, and consequently belief in its existence, was that of the power of transcendence of psychological powers, popularly called second sight. The notion was formulated with a claim for a scientific basis in telepathy or thought-transference. Telepathy includes reading another's mind, seeing distant events, dreaming dreams that come true, communing with the dead, having visions of the time of death, having premonitions of disaster, and much else. To appraise the validity of this body of attested evidence became the major program of societies of psychic research. (The British society was formed in 1882.)

SPIRITUALISM

Emerging from this belief, modern spiritualism began in 1848 with physical manifestations in the presence of two young girls, the Fox sisters, in Rochester, New York. There followed table liftings and rappings which, by the use of a convention that one rap should mean *yes* and two *no,* provided communication. Thus messages were spelled out letter by letter, by yes-and-no raps as the letters of the alphabet were recited—a procedure familiar among the Romans who used for the same purpose a ring striking the side of a glass (in which it was suspended by a thread) when the correct name or answer was spoken.

There was wave upon wave in the revival of interest in psychic transcendence, always brought to a focus in the performance of some medium. Mme. Blavatsky held the stage as a wonder-worker for thirty years, until her death in 1891. She introduced theosophy and established a shrine in India. In evidence of this faith, she performed miracles in astonishing and daring variety, transcending all the laws of matter and mind. She posed them as manifestations of the "Akacic" force, derived from communion with Mahatmas, holy men residing in remote Tibet. She obtained a large following as a prophet of transcendence, but her story is largely a chapter in pseudology.

Of very different caliber was an Italian woman of the peasant class, Eusapia Paladino, who was at the height of her fame in 1910. She aroused international attention with so simple a performance as lifting a light table in a darkened room with her hands and feet supposedly controlled by adjacent sitters. By freeing her left leg, she was able to raise the table on the flange of her sturdy shoe;

by other devices, no more baffling, she produced effects which her followers pronounced supernatural.

ECTOPLASM

Several transcendents appeared in the early nineteenth century. The most spectacular was the demonstrator of ectoplasm—a biological concept proposed by Richet, distinguished physiologist. He attributed to the medium, Eva C. (1880), seated in a dark cabinet, the supreme transcendent power of originating plasm, living tissue, derived from her person. Ectoplasm captivated the popular imagination. A mechanical formulation of the marvel referred it to the action of hypothetical levers and psychic struts emanating from the person of a medium. This was the theory of Dr. Crawford, engineer of Belfast (1920–30), whose clever medium was an Irish girl. His books are sincere exhibits in the treacherous field of psychic transcendence. The interest in ectoplasm was later revived by the varied claims of Marjory of Boston: she demonstrated a psychic thumbprint of a deceased individual; she reversed a chemical balance; she produced mysterious sounds, lights, and perfumes. Her "miracles" found wide acceptance.

Psychic transcendence flourishes in waves. The interest in it is kept alive by the reputation of the few men of science who endorse it. Richet, regarding the phenomena as genuine, proposed a metaphysics, which had three divisions: ectoplasm, telekinesis (moving objects without contact), and cryptesthesia, the perception and communication beyond ordinary senses. To him this was a new science; to the rest of the world it was pseudology.

TELEPATHY

Present inquiry into the possibility of telepathy, e.g., extrasensory perception, is legitimate, though not recognized scientifically as profitable. The most recent attempts to bring the problem within the range of rigid demonstration have failed to satisfy the critiques of the existence of any extrasensory phenomena. The spiritualistic hypothesis itself is clearly beyond testing by scientific observation and the crude evidence offered in support of spiritualism has no possibility of verification.

HYPNOSIS

Claims of psychic transcendence have contributed indirectly to the advance of scientific psychology, notably in two historic instances. One led to the investigation of hypnosis; the other to the extension of knowledge regarding the scope of subconscious processes, and that in two directions: involuntary movements and subconsciously received and elaborated information.

The discovery of hypnosis grew out of a sequel to the curative procedures of Franz Mesmer (1733–1815), dramatically staged around his magnetic *baquet,* which was to induce crises and then draw them out of the patient's system. He posited an animal magnetism, mystically allied to the attractive force of the magnet, resident in animal life, but peculiarly precipitated in and through his person. Mesmerism is pseudology, and despite his redeeming traits in opposing the dogmatic medicos of his time, Mesmer was a quack. But hypnosis is real and its apparent marvels are the results of suggestion in hysterically disposed subjects.

SUBCONSCIOUS MOVEMENTS

The use of the divining rod to find underground water or metals is an ancient and still prevalent example of subconscious movements under intelligent direction. Subconscious movements are the explanation of the rappings and messages of tables in spiritualistic seances. The great English physicist Michael Faraday devised (1854) an instrument to prove this fact, while the great French chemist Michel Chevreul presented the evidence. This demonstrator of involuntary movement was transformed into a *planchette* and later (1880 and 1890) into the famous *ouija board*. This instrument is responsible for much release of psychic or supernatural power of many varieties. Ouija demonstrated the wide prevalence of subconscious or automatic powers in peculiarly disposed but normal individuals. Not by its revelations did it add to the sum of knowledge, but by demonstrating the large play of subconscious processes.

AUTOMATIC AUTHORSHIP

Two English antiquarians published revelations through automatic writing of plans for rebuilding Glastonbury Abbey, which had been written by sixteenth century monks in the Latin of their day. The revelations were so precise in location and measurements that excavations were made and the foundations found as described. Two English women, on visiting Versailles in 1901, saw the gardens as in a vision, with persons, structures, costumes, and decorations as historical records showed they appeared in the days of Marie Antoinette. These are instances of automatic authorship. A remarkable example is contained in the writings and pronouncements of

Patience Worth, a maiden of Elizabethan days, who wove the romances and dramatized the scenes of her daily life in the quaint language of that day. Yet the words came from the pencil and voice of a Mrs. Curran of St. Louis (1913), her normal self being incapable of such improvisations. Mlle Smith, a clerk in a shop in Geneva, enacted scenes located on Mars, including the construction of a Martian language, an unfoldment of an Indian drama, with authentic details of A.D. 1400, and held seances in which she impersonated Marie Antoinette accurately, according to historical records.

All of these exhibitions seem far beyond the range of conscious knowledge and accomplishments. The alternative remains of accepting these truly remarkable records as exhibits of extraordinary subconscious authorship, or admitting that, under circumstances, communion may be held with the past in verifiable detail. As a chapter in pseudology, automatic authorship occupies a unique place.

INDIVIDUAL VAGARIES

Individual vagaries are startling pseudoscientific discoveries, which have a moderate following and then decline in interest. A notable instance was a new force discovered by the chemist, Baron Karl von Reichenbach, which he named Od. Od accounted for auroras and the light of the planets, and for the shifting of the seasons; by this odic sense new chemicals could be analyzed; and ghosts were explained as odic emanations. There was also a plant Od for agricultural purposes, and even a psychic Od. A large bookful of evidence was collected supporting Od solely upon the reactions of "sensitives," many of them hysterical young women under the prepossessions of the deluded discoverer.

Dr. Jaeger, a zoologist, discovered the smell-soul of Duft (1874), which was an emanation perceptible to the sense of smell in properly sensitized individuals. Through his own testing, judged by his measured reaction-times under the influence of selected odors, he concluded that the beneficial Dufts could be retained in the body by wearing clothing of pure wool. Thus woolen clothing became popular, while the weird theory behind it was forgotten. Dr. Buchanan (1843), a pure quack, gave psychometry as the name for the ability to sense the flow of emanations from contact with an object. Thereby he envisages the scene and the association in which the object played a part. A specimen of handwriting, a lock of hair, or a keepsake could reinstate an appropriate incident in the life history of one long deceased. Geological specimens and relics of the past were similarly psychometrized.

CULTS

Cults are systematized beliefs which have a devoted or fanatical following. Among vagaries that became cults, two stand out prominently and give rise to quackery galore. One is physiognomy, an ancient lore which was popularized by Johann Lavater (1741–1801). Laveter published sumptuous volumes illustrating how the physiognomies of famous men (and commonplace ones as well) revealed, down to minute details, their mental and moral characteristics. His following continues to this day. Reading character in handwriting belongs to the same group. But handwriting is a motor expression; graphology offers the possibilties while not the actuality of a scientific interpretation. The greatest pseudological splurge in psychology was the work of Franz Gall (1758–1828), a distinguished anatomist, whose theory, called phrenology, documented by volumes of studies, was that the protuberances of the skull could be charted as an anatomical map of psychological traits and propensities.

The errors in both these cults formed a learned and fairly respectable prepossession. The system was degraded to vulgar quackery in the hands of followers under the universal interest in reading character and telling fortunes by whatever signs appeal to popular fancy.

NUMEROLOGY

Newer in design and application, though not in all its technique, is the peculiar pseudology called numerology, launched by an unknown of the early twentieth century. It assigns to the numbers 1 to 9 a set of fancifully chosen human traits—one group of a constructive or favorable type and the other of a destructive or unfavorable type. Each of the twenty-six letters of the alphabet is given a number value, and the sum of these values in your name assigns you to a type, and consequently, a fortune and fate. Pseudology and quackery combine in numerology in extravagant expression.

EVIDENCE

The logical errors involved in pseudologies illustrate the ready temptations to accept the false in thinking. The apparent trend of evidence suggests a false set of ideas as the hypothesis to account for some phenomenon, and a pseudology is born. Ideas, notions, hypotheses are not of equal validity; they gain their strength from their agreement with the general trend of established knowledge. Pseudologies may appear to have this agreement.

Neither Od or Duft, can be thrown out completely as unscientific, though neither is plausible. But the methods by which Od and the Duft were evidenced showed gross ignorance of the ways of proof, especially in ignoring the subjective factor that created the very results which were offered as proof. Physiognomy was a mildly plausible idea, but rather definitely out of range of science because of knowledge even in those days suggesting that it was highly improbable that human traits were indexed by such specific signs. Then, too, the evidence was not substantial and of the type that read common sense into causes. In phrenology there was a strange ignoring by a good anatomist of the complexities of brain-character relations. The notion of a skull map is wholly unsound and slight brain contours are not indicated on the skull. The relations between skull locations and human traits were chosen by a self-deception, i.e., by reading evidence into the arbitrary variations. Idea and evidence were both wrong in this instance. Palmistry is so wanting in either plausibility of idea or of evidence that it fails of scientific consideration; while the idea of numerology is so fantastic that it falls well within the class of delusions evolved by paranoid minds.

Indeed, there is reason to suspect that some pseudologies result from disarranged minds as well as from false thinking.

READING REFERENCES

Jastrow, Joseph, *The Betrayal of Intelligence; A Preface to Debunking*. New York: Greenberg, Publisher, 1938.

———, *Fact and Fable in Psychology*. Boston: Houghton Mifflin Company, 1900.

———, *Wish and Wisdom*. New York: Appleton-Century-Crofts, Inc., 1935.

Mackay, Charles, *Extraordinary Popular Delusions*. 1849.

CHAPTER II

INDIVIDUAL EFFICIENCY

8. Reading Efficiency	William S. Gray
9. Typography and Legibility in Reading	Miles A. Tinker and Donald G. Paterson
10. Illumination	Dean Farnsworth
11. Ventilation	Kinsley R. Smith
12. Fatigue and Its Alleviation	Raymond A. Katzell
13. The Process Conception of Psychological Illness	Matthew N. Chappell
14. Drugs and Smoking	St. Clair A. Switzer

8. READING EFFICIENCY

WILLIAM S. GRAY, PH.D.

Professor of Education and Editor, *Proceedings of the Annual Conference on Reading,* University of Chicago. Formerly: Dean of the College of Education, University of Chicago and Executive Secretary of the University's Committee on the Preparation of Teachers; President, the American Education Research Association, Section Q of the American Association for the Advancement of Science, and the National Society of College Teachers of Education; Chairman, National Committees on Reading and Preparing Reports for the National Society for the Study of Education.

The deep concern today on the part of individuals about their reading efficiency has its origin in several significant facts. The amount of reading material published has increased so rapidly during the last decade that young people and adults encounter difficulty in reading even those portions of it that relate directly to their interests and needs. As a result the question often arises, How can I increase my speed of reading so that I can read more widely in the time at my disposal? Since the books and articles read usually relate to a wide range of topics or themes, and often include unusual terms and unfamiliar concepts, the reader may encounter serious difficulty in comprehending the ideas presented. As a result he often asks, What can I do to improve my efficiency in securing a clear grasp of the meaning of what I read? It is not easy for individuals to answer such questions satisfactorily when the answers involve a psychological analysis of reading or suggestions for improving ability to read.

NATURE AND PURPOSES OF READING

Of major importance is the fact that the concept of reading has expanded rapidly during recent years. Whereas reading was formerly defined largely in terms of the basic attitudes, habits, and skills involved, it is now conceived primarily as a form of experience, a mode of learning, and a way of enjoying leisure time. A significant measure, therefore, of reading efficiency is the extent to which it enlarges the readers' experiences, contributes insights and understandings, broadens interests, cultivates attitudes and ideals, modifies behavior, and promotes rich and stable personalities. Thus conceived, reading differs little from other forms of experience such as listening to the radio, attending a movie, or participating in a discussion. In each of these cases stimuli are received, and if the experience is genuinely effective, meanings are apprehended, ideas are interpreted, and some form of reaction occurs.

READING PURPOSES

Equally significant is the fact that reading serves a wide range of useful purposes. Young people and adults read silently to keep informed about current events, to secure needed information, to learn about a given topic, to find out an author's views, to solve a personal problem, to decide how to act in new situations, to reach conclusions concerning challenging social issues, to interpret human life and motives, to satisfy natural and valuable curiosities, to find escape and relaxation, and to derive pleasure. They read orally to support a view or defend an argument, to provide information, instructions, or directions, to entertain others, and to enjoy thoroughly the rhythm and beauty of poetry. The situations in which oral reading occurs varies from the formal demands of a radio or other form of public presentation to those of the intimate family circle where reading aloud contributes enjoyment to all or arouses interest in reading on the part of children. Although the foregoing list is far from complete, it indicates that both silent and oral reading meet vital needs today.

THE GOOD READER

Certain individuals are able to read effectively only a few kinds of material or for only a limited number of purposes. To be a good reader in the broadest sense of the term one should be able to engage effectively in the various types of reading activities that are essential in contemporary life. This makes a heavy demand on the individual as indicated by the analysis that follows.

PSYCHOLOGICAL PROCESSES

When one reads a sentence, paragraph, or longer selection properly a large number of processes are involved. The eyes move from left to right along the lines; meanings and pronunciations are recognized; and the thoughts and ideas presented are clearly comprehended and interpreted. Because of their very nature many of the steps and processes that occur cannot be observed directly. It is possible, however, to identify many of them through studies of the behavior of children and adults while reading and of the steps involved in related mental activities.

BASIC PROCESSES

The basic processes in mature reading have been described by psychologists and are summarized here in the following statements:

1. When a person begins to read the retina of the eye receives stimuli or impressions from the printed page. As the eye moves forward along the line the retina receives a series of stimuli.
2. The impressions received on the retina arouse nerve impulses that pass to the visual centers of the brain.
3. As impulses from the retina reach the appropriate brain centers associations are aroused that result in the recognition of meanings and pronunciations and in the comprehension and interpretation of the ideas represented by the printed symbols.
4. As recognition, comprehension, and interpretation occur impulses pass to the motor speech centers. This is particularly true in the case of oral reading. They occur also in good silent reading, but far less prominently.
5. Impulses pass from the motor speech centers to the muscles of the vocal cords, the lips, the tongue, and related parts.
6. The final step in oral reading is the appropriate movement of the speech organs in expressing the ideas apprehended. In the case of silent reading they are often so suppressed that they involve only incipient movements of the throat.

In order to describe more fully what is involved in efficient reading the foregoing outline of basic processes will be supplemented by more detailed discussions of various aspects of the reading process.

TYPES OF ASSOCIATIONS AROUSED

When the impulses from the retina reach the appropriate brain centers several types of associations are aroused. The most important in both oral and silent reading are those that aid in the recognition of meaning.

Meaning Associations. The fullness and richness of the meaning associations that are aroused depend upon the experiences of the reader in the specific fields to which the passages relate and on the habits of thinking while reading which the reader has acquired. A wide meaning vocabulary is, of course, an asset of great importance. It correlates more highly with efficient silent reading than any other factor except the intelligence of the reader. Experiments show that vigorous effort on the part of individuals to extend and clarify their understanding of the meanings of words results, as a rule, in improved comprehension. Methods used include the study of the context in which a word is used, the identification of roots, prefixes, and suffixes as clues to meaning, and the discriminating use of a dictionary.

Pronunciations. A second type of association that is aroused aids in recognizing the pronunciation of words. In the case of the good reader these associations occur more or less unconsciously, which is due to the fact that the sight and sound of words have been associated with sufficient frequency so that stimuli from printed symbols arouse at once associations relating to pronunciation. For the poor reader, however, the recognition of pronunciations may occupy the center of attention. In such cases the reader is often required to analyze a word before he is able to pronounce it. The following aids to pronunciation are particularly valuable: the meaning or context of the passage in which the word occurs, the general form or configuration of a word, significant or distinguishing details, structural parts of words, phonetic elements, and the dictionary. A good reader utilizes one or more of these aids as they serve his needs best in reading a given passage. A poor reader, on the other hand, is unable to use them effectively and consequently encounters frequent and insurmountable difficulties.

Oral Interpretation. A third type of association relates to the methods, devices, and principles of oral interpretation that underlie effective reading to inform or entertain others. The interests of the audience, the message to be transmitted, the mood of the selection, the volume and modulation of the voice, and the most effective ways of expressing the author's meaning must be considered. The reality of the associations that influence oral interpretation is very clear when a trained reader studies a selection in preparation for an audience presentation.

COMPREHENSION OF WHAT IS READ

Far more is involved in a clear grasp of the meaning of a passage than the arousal of the various associations described. The additional steps in comprehending what is read can be described by discussing three different aspects of meaning, namely, the literal meaning of a passage; the meaning implied by the author's mood, tone, and purpose; and the additional meanings that arise out of the specific purposes and unique experiences of the reader. The apprehension of meaning is a unitary process, but its different aspects will be discussed separately in order to make clear various characteristics of efficient reading.

LITERAL MEANING. The literal or sense meaning of a passage is what is implied when the reader is asked, "What does the sentence or paragraph say, mean, or tell you?" In order to answer such a question the reader recalls or discovers the meanings of the different words and phrases in the passage. Frequently he must search actively for the meaning of given words as implied by the context. As appropriate associations are aroused he fuses or organizes them into a chain of related ideas. This results not only in a grasp of the sequence of meaning but also in the recognition of the relative importance of the ideas and their relationship to each other.

The heart of this part of the reading act is to select and combine the various items of meaning into a consistent whole. Good thinking is required at every step in apprehending meaning. As pointed out by Thorndike (1917), the reader must select, repress, emphasize, relate, and organize as he attempts to secure a clear grasp of the meaning of what he reads.

A reader not only uses past experience in associating meanings with words, but also may acquire new and vivid experiences through reading about things that lie beyond his immediate environment. Thus one reads in a geography of a land that is always covered with ice. The word 'ice' is familiar and the word 'always' also has acquired definite meaning. When the two ideas are properly related the reader is carried into a world which is unknown, but may be understood through the use of meanings acquired in everyday life. Reading is thus an active creative process which requires the undivided attention of the reader. The fact that reading is a significant means of extending and enriching experience makes it of great value at every level of education, both formal and informal.

AUTHOR'S PURPOSE AND MOOD. Good reading includes both a clear grasp of the literal meaning of a passage, and in addition the meaning implied by the author's mood, tone, and purpose (Richards, 1938). Indeed, a reader must make use of all that throws light on the ideas represented by the printed words. This often involves a knowledge of such items as the time and place in which a passage was written, the setting in which the described events occurred, the general characteristics and interests of the author, and his mood, tone, and purpose as revealed by the choice of words used and the arrangement of the ideas presented.

Good reading in this broader sense is hard to attain. Many readers fail to apprehend fully and clearly because they do not have an adequate background of related experience; others fail to recall and make use of relevant facts which they know. Of major importance is an inquiring mind which leads the reader to seek diligently for the author's meaning, often using other aids to learning in acquiring sufficient background to read given types of material effectively.

READER'S PURPOSE. A reader sometimes attaches meanings to a passage which the author did not have in mind when writing it. An orator on Decoration Day, for example, may attach meanings to the words of Lincoln's Gettysburg Address which its author could not have meant when he delivered it. This is due to the particular experiences, interests, and biases of the orator.

Because the reader may make such interpretations, it is imperative that he distinguish clearly between the meanings intended by the author and his own interpretations of them. The latter should be checked carefully and should be given no more consideration than they merit. Various reading specialists have pointed out recently that one of the most serious deficiencies in the reading of both youth and adults is failure to distinguish clearly between the author's intended meaning and what they individually read into a passage.

The reader's immediate interests modify materially the direction of his attention while reading, the thoroughness with which specific meanings are grasped, and the value that he attaches to each. At times he may wish to secure a clear and relatively complete grasp of the author's meaning; on other occasions he may be concerned only with those items of information for which he has immediate need. Similarly, the meanings to which major attention is directed may be determined by the kinds of material read, for example, a news item, an editorial, the President's message to Congress, a detective story, directions for filling out income tax reports.

REACTING TO AND APPLYING WHAT IS READ

Valuable as they are, the recognition of symbols and the comprehension of meaning do not comprise the whole of the reading act. In addition, the reader reacts critically to what he reads, applies the ideas gained in the solution of problems, and fuses what he reads with previous experience so that new understandings and rational attitudes result.

CRITICAL READING. The efficient reader reacts critically in many ways to what he reads. For example, he judges the worth or validity of the ideas presented and the quality or completeness of the author's presentation. He identifies the basic assumptions underlying the arguments set forth and accepts or rejects the author's conclusions depending on their validity or appropriateness. He reacts to the literary quality of what he reads and is pleased or disappointed. The extent and character of the critical thinking that accompanies the reading act depends in large measure on the reader's purpose. It is clear, for example, that critical thinking may be very prominent when reading an article to determine the validity of the author's views on a topic; it may be very slight when reading a short story for pleasure and relaxation.

The habit of critical reading has broad personal and social implications. It insures a growing body of concepts and ideas, the validity and relationships of which are clearly recognized; it aids in the development of attitudes, ideals, and appreciations that are the product of discriminative thinking; and it enables a person to change advisedly his outlook on life, to modify his purposes and intentions, and to chart new courses of action with care and discrimination (Gray, 1940). One cannot read critically as rapidly as he does when merely apprehending the sense meaning of passages. However, the content of what is read has greater value to him because its validity and significance have been considered. Among the requisites of critical reading are an inquiring mind, an adequate background of information, ideas, and principles to serve as a basis of judgment, and willingness to put forth the effort necessary to engage in critical and discriminating thinking.

APPLYING WHAT IS READ. The efficient reader also applies constantly the information and ideas secured from the printed page. This step is essential if he is to be aided through reading in solving personal and social problems and in acquiring increased power of self-direction. It involves the integration into definite thought and action patterns of pertinent concepts and ideas gained through reading and from previous experience. This results in the clarification of thinking, in improved judgments, and in better guides to action. Many individuals read widely without making significant use of what they read. The good reader recognizes the implications and relationships of what he reads and combines the ideas apprehended with previous experience in such ways that new insights and understandings are acquired, challenging problems are solved and valuable guides to behavior are developed.

MODIFYING PERSONALITY. The foregoing discussion implies that reading changes personality. As readers comprehend accurately, interpret broadly, react critically, and apply what they learn, they acquire new understandings, broader interests, and deeper appreciations. Personality is thus continuously modified through reading. The fact is recognized that reading, as well as other forms of experience, may exert either a negative or a positive influence on the reader. A well balanced diet of reading, accompanied by critical interpretation, should reduce negative influences to a minimum and should aid in developing a generation of citizens with rich and stable personalities.

SPEED OF READING

One of the striking facts revealed through studies of reading (Vernon, 1931) relates to wide variations in the speed at which different people read the same material. Similarly, a given reader varies widely in speed when reading different kinds of material or for different purposes. On some occasions the purpose of the reader can be attained best through very rapid reading, for example, scanning an article quickly to find out if it contains material relating to a specific problem. On other occasions, he reads slowly to secure a clear understanding of an author's point of view, to determine the implications of an article, or to clarify his thinking concerning an issue discussed. Of major importance is the fact that the good reader is conscious of the purpose to be achieved through reading on given occasions and reads only as rapidly as he can achieve his purpose well. All other things being equal, however, the more rapidly a reader can achieve specific purposes, the greater his efficiency in reading.

FACTORS DETERMINING SPEED. The fact has long been recognized by psychologists that speed of reading depends far more on the facility with which the mind can interpret impressions than on the speed with which impressions are secured. In harmony with this view, the results of scientific studies show clearly that a relatively high level of intelligence, familiarity with the vocabulary and theme of a passage, and ability to grasp ideas quickly are among the most important factors influencing speed of reading. They show also that accuracy of recognition, the amount read at each fixation of the eyes, and the regularity of the progress of perceptions from left to right along the lines influence speed of reading. The efficient reader, therefore, is one who proceeds rapidly in the recognition and interpretation of meaning and who adjusts his speed to the intellectual effort necessary to achieve the purpose that takes him to the printed page.

INCREASING READING EFFICIENCY

Many techniques have been found through experiments or careful observations to have distinct value in improving reading efficiency. None of them, however, has the miraculous quality of converting a poor reader into an efficient reader in a short period of time or without vigorous effort on his part. But, if properly applied, they usually improve reading efficiency.

READING PURPOSE

The first and most important step in improving one's efficiency in reading is to approach any reading activity with an inquiring mind or with a purpose or problem that is significant or worth while. Keen interest in the content of what is read and a strong motive for reading a given selection have been shown repeatedly to be of great value in increasing reading efficiency. To have a purpose for reading provides the mental set and inner drive essential in directing attention effectively to the content of what is read and in maintaining sustained effort.

A brief period of mental preparation for reading is often highly desirable. A British man of letters was once asked how he was able to read so much and so well during a given evening. In reply he stated that he first prepared his mind for reading a given book by asking himself a series of questions such as the following: What is the probable theme of this book as suggested by the title? What would be the most timely thing to say on such a theme today? Who wrote the book? What kind of writer is he—constructive or cynical? What will his views probably be on this theme? What are my own views concerning it? The Britisher reported that with such a preparation he was able to read a book rapidly and with absorbed attention to the content. He was able to read critically because in the preparatory step he had identified significant issues relating to the theme. It is not assumed here that the specific steps outlined are appropriate for all types of reading. The important fact meriting emphasis is that efficiency in reading is greatly increased by a brief mental survey of the possible content, implications, issues, and contributions of the materials to be read.

MIND-WANDERING

Reading efficiency is greatly reduced if an individual engages in a great deal of mind-wandering while reading or is unable to keep his attention on the content of the passage. Failure to read with absorbed attention usually is due to lack of genuine interest, motive, or drive. When a reader finds his mind wandering, he should recall his purpose for reading or the nature of the contribution he seeks. If competing interests are very persistent, he should either suppress them or yield to them completely for the time being. To continue reading while the mind is primarily engaged with other matters is futile.

THINKING WHILE READING

Closely associated with mind-wandering is failure on the part of a student or adult to carry on relevant and productive thinking while reading, that is, to distinguish between the important and the unimportant, to grasp relationships between the ideas apprehended, to identify strength or weakness in the author's views, to recognize propaganda, to reach valid conclusions about an issue, and to integrate the ideas apprehended and previous experience into valuable thought and behavior patterns.

Efficient reading in this respect is the product of vigorous and persistent effort. To establish such habits, some psychologists request students to engage under their direction in reading which is preceded by stimulating questions and interrupted after each short unit with discussion which aims to check their understanding of what they read. As soon as the students acquire a definite notion of the kind of thinking essential in the interpretation of assigned materials they are required to assume responsibility for directing their own thought processes in further reading activities. Experience teaches that most students can greatly increase their competence in reading by raising questions constantly as they read, and by seeking to discover significant relationships and implications.

MEANING VOCABULARY

A good reader is highly sensitive to the meanings of specific words and endeavors constantly to enlarge his meaning vocabulary. A valuable first step is to study the meanings of words as indicated by their context. Not infrequently the meanings implied are different from those previously attached to the same words by the reader. In such cases the study of words in context, not only increases the accuracy of interpretation, but enlarges the meanings of familiar words and attaches meanings to new words.

Experience shows that the reader who makes mental or written note of such words grows rapidly in his grasp of meanings and in the interpretation of what is read. Likewise readers who study the meanings of roots, prefixes, and suffixes and use them as aids in securing the meanings of words usually grow rapidly in reading efficiency. Furthermore, the importance of the frequent but intelligent use of the dictionary while reading cannot be overemphasized.

READING SPEED

The most frequent request for help in improving reading that is made by adults relates to speed of reading, which is usually due to the fact that they cannot do all the reading that is required or that they would like to do in the time available. Furthermore, some have never learned to read fluently and proceed along the lines at an unusually slow rate. Some of the conditions that favor rapid reading are keen interest in the content of a passage, ability to read with undivided attention, a relatively wide background of related information, familiarity with the meaning and pronunciation of the words used, and purposes for reading that do not make heavy demands on the reader for critical evaluation and judgments.

Many procedures have been adopted to increase speed of reading. They include the reading of simple and highly interesting material for periods of a half-hour, or more, daily to establish fluent habits of reading; conscious effort to increase speed of reading by keeping a record of the amount read in a given period of time and endeavoring to improve that record during the next reading period; using a metronome as an aid in establishing rhythmical eye movements in reading; and reading at times under controlled conditions, such as those provided through the use of reading films (The Harvard Reading Films, Harvard University, Cambridge, Mass.), the metronoscope (The American Optical Company, Southbridge, Mass.) and reading boards (The Three Dimension Company, 900 North Dearborn Street, Chicago, Ill.).

As a result of experiments involving the use of the foregoing methods certain conclusions have been reached which merit emphasis. The first is that speed of reading can be improved without decreasing comprehension, provided the reader holds himself responsible for a clear grasp of meaning. The second is that in the case of very immature readers definite training in basic habits of word recognition and comprehension is necessary before speed can be materially increased. The third is that whereas mechanical aids are definitely helpful in increasing speed of reading among some types of slow readers, such as the so-called "word reader," definite and permanent improvement can be made without them.

READING REFERENCES

Adler, Mortimer J., *How to Read a Book*. New York: Simon and Schuster, Inc., 1940.

Buswell, G. T., "How Adults Read." *Supplementary Educational Monographs,* 1937, No. 45.

Chase, Stuart, *The Tyranny of Words*. New York: Harcourt, Brace and Company, Inc., 1938.

Gray, W. S. (ed.), *Reading in General Education*. Washington, D. C.: American Council on Education, 1940.

Hamilton, J. A., *Toward Proficient Reading*. Claremont: Saunders Press, 1939.

Lee, Irving J., *Language Habits in Human Affairs*. New York: Harper & Brothers, 1941.

Progressive Education Association, Commission on the Secondary School Curriculum, *Language in General Education*. New York: D. Appleton-Century Co., 1940.

Richards, I. A., *Interpretation in Teaching*. New York: Harcourt, Brace and Company, Inc., 1938.

Starch, D., *Educational Psychology* (rev. ed.), Chap. 16. New York: The Macmillan Company, 1927.

Thorndike, E. L., "Reading as Reasoning: A Study of Mistakes in Paragraph Reading." *Journal of Educational Psychology,* 1917, **8,** 329.

Vernon, M. D., *The Experimental Study of Reading*. New York: The Macmillan Company, 1931.

Westover, F. L., "Controlled Eye Movements versus Practice Exercises in Reading: A Comparison of Methods of Improving Reading Speed and Comprehension of College Freshmen." Teachers College, Columbia University, *Contributions to Education,* 1946, No. 917.

Woodworth, R. S., *Experimental Psychology*. New York: Henry Holt and Company, Inc., 1938, Chaps. 23 and 28.

9. TYPOGRAPHY AND LEGIBILITY IN READING

MILES A. TINKER, PH.D.

Professor of Psychology and Director, Lighting and Reading Laboratory, University of Minnesota; Consultant on Problems of Illumination, Typography, and Reading Efficiency.

DONALD G. PATERSON, M.A.

Professor of Psychology and Member of Staff, Industrial Relations Center, University of Minnesota; Editor of the *Journal of Applied Psychology*. Formerly: Psychological Officer, US Army; Personnel Consultant, The Scott Company; President, The American Association for Applied Psychology.

Few people realize the extent to which reading dominates contemporary life. It is well to remember that learning to read is the most important thing accomplished during the first few years in school. Thereafter reading becomes the primary tool for learning all other school subjects. It becomes more important in high school and college where an even heavier demand is placed on the ability to read rapidly the increasingly difficult material.

In all the reading one does in school, at work, or in his leisure time, it is clear that constant pressure is placed upon the individual to assimilate a maximum amount of printed material in a minimum amount of time. Furthermore, few individuals can meet these demands without struggling to increase the efficiency with which they read. A basis for the improvement of reading conditions is furnished by an analysis of the reading act. In addition, optimal typographical conditions promoting ease and comfort in reading must be discovered.

ANALYSIS OF READING ACT

The analysis of the reading act has occupied the attention of many psychologists since the days of Huey (1909). The primary approaches are two: perception and eye-movement studies, and measurements of intelligence and reading achievement.

In general, two groups of factors have been isolated, namely, peripheral and central. Under peripheral would be listed the physiological condition of the visual mechanism (visual acuity, muscle balance, etc.), as well as the intricate and complex eye movements involved in different kinds of reading. The central factors refer to the role of intelligence, range of vocabulary, and the ability to see meaningful relationships between words, phrases, and larger units.

The physiological condition of the eyes does not necessarily determine the degree of efficiency with which one reads. Nevertheless, whenever deficiencies are found, corrective exercises and optical aids should be provided. In this way visual fatigue and feelings of eyestrain may be overcome with a consequent improvement in the physical welfare of the individual. The optician, the optometrist, and the ophthalmologist make their main contribution in this area.

Photographic methods have been used to provide an objective basis for analyzing oculomotor patterns in reading. These patterns show that the eyes, when one reads a printed line, move in a series of jumps separated by a series of pauses (Vernon, 1931). Clear vision exists only during the pauses so that reading involves the fusion of successive visual perceptions. A good reader differs from a poor reader in the number, duration, and regularity in the sequence of the pauses in reading a given line of print.

Eye-movement photography also shows that reading does not ordinarily involve discrimination of small details, such as individual letters. Words and short phrases are the units of perception for the efficient reader. These are the "thought units" which underlie effective reading. This finding leads to the conclusion that peripheral factors, for example, eye movements, primarily reflect central factors. In other words, an eye-movement pattern is merely a symptom of the efficiency with which the central processes operate.

The importance of intelligence as a determinant of reading proficiency is revealed by the method of correlation. Strang and Rose (1938), in summarizing studies of the correlation between standard intelligence tests and standard tests of reading proficiency, reports that the majority of coefficients range between 0.50 and 0.70.

That range of vocabulary is considered to be intimately related to reading ability is reflected by the fact that standard reading tests are likely to contain a direct or an indirect measurement of vocabulary. Indirect measurement is achieved by means of comprehension tests in which paragraphs of increasing difficulty are provided. Increases in difficulty are brought about in part by the use of more and more difficult words.

TYPOGRAPHY AND LEGIBILITY

In addition to the factors cited above, one must recognize the importance of those physical aspects of the environment which condition reading efficiency. Here one is concerned with suitable illumination and optimal typographical arrangements. The problem of illumination is discussed in other articles and the typographical factors will be treated in the present article. In what follows the terms legibility and readability are synonymous. Likewise, the terms perceptibility and visibility are synonyms.

Two reading situations must be kept in mind. The first is concerned with reading printed matter at the ordinary or usual distance of approximately 14 inches. Here we deal with the legibility of print in newspapers, magazines, books, pamphlets, circulars, etc. The second involves the reading of printed matter at a distance. Here one is concerned with the perceptibility of newspaper headlines, billboards, car cards, window display signs, road signs, labels on packaged goods, etc.

The two kinds of reading situations described require that a distinction be made between readability or legibility (speed and ease of reading at the normal distance with adequate illumination) and perceptibility (visibility at various distances or visibility at the normal distance with inadequate illumination). The distinction is more than verbal since typographical arrangements that can be read accurately at various distances or under conditions of poor illumination may not be suitable for rapid and easy reading at the normal reading distance with adequate illumination.

LEGIBILITY

The problem of measuring the relative effectiveness of various printing arrangements is by no means simple. Complications arise because one printing arrangement may be read more rapidly than another and yet may involve eyestrain and visual fatigue to such an extent as to make its use inadvisable. Suitable speed of reading tests are available but suitable fatigue tests are not. It seems advisable at the present time, therefore, to use speed tests accompanied by reader judgments of legibility. The purpose of these two approaches is to ascertain the best typography to insure the selection of a printed page that may be read rapidly without feelings of strain.

Another method of measurement is provided by photographing eye movements during reading. This method is far more analytical than speed of reading tests since it permits precise analysis of oculomotor patterns as revealed by number of fixational pauses per line, duration of pauses, and retracings or regressions. The disadvantage of this method is chiefly one of expense in time and money. This fact ordinarily precludes using adequate amounts of reading material and a sufficiently large number of readers to yield reliable comparisons. Nevertheless, it is of great value as a supplement to speed tests. Such tests, given to large numbers of readers can be used, quickly and inexpensively, to determine which of several typographical arrangements is best. Eye-movement photography can then be employed to analyze the reasons for such superiority.

Luckiesh and Moss (1942) have proposed the blink technique as a means of studying the legibility of type. This technique simply requires counting the number of eyelid blinks made in reading a printed text for a set period of time. The fewer the blinks, the better the printing. It is possible that this method reflects ease of reading but does not necessarily yield an adequate measure of speed of reading. It is quite likely therefore, that eyelid blink frequency is more closely related to visibility of print than it is to legibility.

VISIBILITY

Visibility, or perceptibility of print, refers to the reading of print at the threshold for visual discrimination of words. The threshold itself may be determined by finding the least amount of illumination necessary to perceive the printed words at the usual reading distance or by measuring, under conditions of adequate lighting, the distance at which words can be perceived.

A visibility meter has been invented by Luckiesh and Moss (1942) to determine the threshold of perceptibility at the usual reading distance. In addition, the distance method, used for many years in psychological laboratories, enables one to measure under constant but adequate illumination the distance at which any particular bit of printing can be perceived. Presumably both methods measure the same thing. In other words, type arrangement A, if found to be better than B by one method, would also be found to be better by the other. Evidence also points to the fact that both techniques are appropriate only for determining the best printing arrangement to be employed under unusual conditions, namely, reading at the ordinary distance in a very faint light, or where there is poor contrast between printed words and background, or when reading at a considerable distance. Furthermore, there is evidence that although arrangement A may be better than B in terms of visibility or perceptibility, yet B may actually be read more rapidly than A under ordinary conditions of reading.

There is also available the tachistoscopic or short exposure method for measuring perceptibility of type. This method is especially useful for determining the accuracy of perceiving isolated words, short phrases, and sentences in a momentary exposure. It is clear that it has direct application to problems of the best typographical arrangement for newspaper headlines, advertising slogans, road signs, license plates, and other kinds of printing in which reading at a glance is essential.

The problem of typography is complicated both by the fact that legibility and visibility are somewhat independent, and by the large number of typographical factors involved in printing. We refer to such variables as type form, style of type face, size of type, line width, interlinear spacing, color of ink and background, paper surface, and effect of margins. Space precludes adequate discussion of each of these factors in relationship to the others. Paterson and Tinker (1940) give detailed evidence on these problems.

A TYPICAL PROBLEM

In order to show how the applied psychologist solves a typographical problem, attempts to determine the relative legibility and perceptibility of printing in all-capitals and in lower case will be described. This is a special problem of direct interest to those who write printing specifications for headlines and other brief messages in advertisments, such as posters, billboard advertising, circulars, as well as names and titles on books, packages, and calling cards. One need only look at several advertisements or sub-headings in chapters of books and he will note that designers do not agree as to the relative legibility of these two kinds of printing since both kinds are so freely used. If there were a marked difference, a precise determination of this difference would point the way toward the consistent use of the more legible kind of copy.

In order to determine the relative legibility of all-capitals and lower case printing in the ordinary reading situation the writers presented to 320 readers two equivalent forms of a reliable speed of reading test. One form of the test was set in all-capitals and the other form in capitals and lower case. The results showed a clear-cut difference in favor of lower case. The all-capitals text retarded speed of reading by about 12 per cent. Few typographical factors will retard reading to this extent.

Photographs of the movements of the eye employed in reading all-capital text and in reading lower case were also made. The results show that the difference between all-capitals and lower case is due to a marked increase in the number of fixations or pauses required to read a given amount of printed material. This in turn appears to be due in part to the fact that the all-capital text requires a 35 per cent increase in the amount of printing surface. In addition "word form" cues are lacking.

The importance of word form as an aid to perception may be illustrated by printing in parallel columns the same words in caps and lower case as shown below:

STOPPED	Stopped
YESTERDAY	Yesterday
ALL-CAPITALS	All-capitals
APPLES	Apples

The same contrast is shown by the following headline printed both ways:

BABY FIGHTS FOR LIFE
Baby Fights for Life

If one makes a block outline of the words printed above, it will be obvious that each of the block outlines for the words printed in all-capitals will be rectangular in shape, whereas each of the outlines for the words printed in capitals and lower case will be irregular in shape. It is the lower-case letters with ascenders and descenders that produce the more characteristic "word forms."

In addition to objectively determined legibility, one should consider the opinions and prejudices of readers with respect to what they believe is most legible. To check on this point we submitted samples of all-capital printing and lower case printing to 228 readers with the request that each indicate which kind of printing he believes would be read faster and with the greater ease. Again the results were clearcut—90 per cent stating that lower case is more legible.

The above facts provide a sound basis for improving the printed page. Reliance should be placed on caps and lower case, and the use of all-capitals should be avoided. But some may believe that all-capital printing of words and short phrases is aesthetically superior to lower case. We submitted samples of all-capital printing and lower-case printing, therefore, to another group of 140 readers with instructions to state which kind of printing is more pleasing. The result was decisive. Ninety-four per cent voted in favor of lower-case printing.

The distinction between legibility at the usual reading distance and visibility at a distance, mentioned earlier in this chapter, is emphasized by a study of the visibility of printing in all-capitals and in lower case. Here the distance method shows that words printed in all-capitals can be recognized at a much greater distance (39 per cent greater) than the same words printed in lower case. If one were concerned only with sheer perceptibility then he would advocate the use of all-capitals in those situations where a message must be read at a considerable distance. Unfortunately, such situations are the very ones which require emphasis on speed as well as on perceptibility. For example, billboard advertising and newspaper headlines depend on getting the message across in a fleeting glance. If the distance of the message from the reader is such as to make a given size of printed type barely perceptible, then it would be desirable to increase the size of type and to use lower-case printing rather than all-capitals in order to secure the advantage of quick perception as well. Ample evidence on this problem is provided by Paterson and Tinker (1946).

A consideration of the facts regarding the legibility, visibility and aesthetic value of lower-case printing versus all-capital printing leads to the obvious recommendation that all-capital printing should be abandoned. Reliance should be placed upon lower case, using larger type sizes if necessary. The visibility of lower case can also be increased by specifying the use of bold face type. In these ways maximum effectiveness will be achieved from every point of view.

There remains, however, a use for all-capital printing. We refer to the fact that an occasional use of all-capitals can be justified as a mechanical device for attracting attention. For example, if one looks at a page of want-ads he will notice that words and phrases printed in all-capitals stand out on the printed page. This is due to the attention-getting value of infrequently appearing words or phrases set in all-capitals. This increased attention-value, of course, arises from novelty and contrast.

RECOMMENDED TYPOGRAPHY

In a similar manner the search for optimal printing arrangements with respect to a large number of typographical factors has been completed by Paterson and Tinker (1940). The findings may be summarized as follows:

STYLE OF TYPE FACE

Any commonly used type face or ultramodern type face can be used since they are equally legible. Readers, however, prefer a type face which borders on the bold face. Typewriter type and Old English should be avoided in ordinary printing because of poor legibility. Typewriter type, however, might well be used when a novelty effect is desired. Old English might also be used when brief messages of a sacred, solemn, or formal kind are involved.

TYPE FORM

(*a*) Italics are read more slowly than lower case Roman text. Readers overwhelmingly choose Roman as being more legible as well as more pleasing. *For these reasons italics should be employed sparingly and then only when added emphasis is desired.* (*b*) Bold face printing is read just as rapidly as ordinary lower case. Readers believe, however, that bold face is less legible and less pleasing. **Bold face, therefore, should be used only when emphasis is desired or when perceptibility at a distance is involved.** (*c*) All-capital printing is read much more slowly than lower case. READERS JUDGE ALL-CAPITALS TO BE QUITE ILLEGIBLE AND TO BE FAR LESS PLEASING THAN LOWER CASE. All-capital printing, therefore, should be avoided except for those rare occasions when it may be used as a device to attract attention.

SIZE OF TYPE

Small sizes of type (6 and 7 point) and unusually large type sizes (14 point or larger) are less legible and should be avoided. The intermediate type sizes

(9, 10, 11 and 12 point), when appropriate space between lines and suitable line widths are used, are all equally legible and should be relied upon for ordinary printing. Of these sizes, 11 point is liked slightly more by readers. Larger type sizes should be used, of course, where emphasis is desired or where perceptibility at a distance is involved.

WIDTH OF LINE

Excessively short line widths (less than 2⅓ inches) and excessively long line widths (more than 4⅔ inches) are difficult to read and should be avoided. Intermediate line widths (between 2⅓ and 4⅔ inches) are more legible than the shorter and longer lines. Readers prefer moderate line widths because they believe them to be more legible and more pleasing. In general, therefore, line widths in the neighborhood of 3 inches are recommended.

INTERLINEAR SPACING OR LEADING

Leading (2 point) improves the legibility of all type sizes except 12 point and larger. Readers likewise prefer 2-point leading from the point of view of legibility as well as on the basis of aesthetics. In general leading is found to be more important than line width.

TYPE SIZE, LINE WIDTH AND LEADING

Three factors, type size, line width, and leading, are intimately interrelated and should be considered together in the writing of printing specifications. To take into account optimal amounts of leading and line widths for each size of type, a series of tables showing safety zones is presented in the Paterson-Tinker book (1940).

MARGINS

Most book and magazine printing devotes far more space to margins than is necessary. In fact the average book wastes a considerable part of each page since half of each page consists of marginal space. Legibility studies show that margins do not promote fast reading. It is probable that in the future an increased proportion of each page will be devoted to print.

COLUMNAR ARRANGEMENT

The trend toward multiple-column printing in magazines, scientific journals, and reference books will probably be extended more to book-printing, where it is now employed occasionally. Double-column printing is preferred by readers over single-column printing. Double-column printing would introduce desirable economies in printing. At the same time it would promote legible printing by preventing the frequent use of excessively long lines of print. It would be unfortunate, however, if the adoption of double-column printing should lead to the use of excessively short lines.

SPACE BETWEEN COLUMNS

In multiple-column printing it is found that an intercolumnar rule or line does not promote legibility in comparison with the use of a one-half pica space (about $\frac{3}{32}$ inch).

COLOR OF PRINT AND BACKGROUND

(*a*) For ordinary printing as well as for material to be read at a distance, black print on a white background is much more legible than white on black. There is an obvious application of this finding to the preparation of advertisements, billboards, posters, car cards, signs, license plates, instrument dials, keyboard labels, etc. The occasional use of white on black for brief messages can be justified, however, as an attention-getting device. (*b*) Colored print on colored backgrounds is less legible than black on white. If used, the law of brightness contrast should be observed. This means that a dark-colored print should be used on a light-colored background. A dark print on a dark background or a light print on a light background should be avoided.

PAPER SURFACE

Printing on highly glazed paper stock is just as legible as printing on dull paper stock. Nevertheless, because of the prevalence of unsatisfactory conditions of illumination, one should employ as far as possible dull finish, opaque paper stock.

GENERAL CONCLUSION

Studies concerned with the legibility of type point to an amazing adaptability of the human eye in adjusting to a bewildering variety of typographical arrangements. We readily turn from reading large headlines to reading news stories set in small type and short line widths, and then to larger type sizes and longer line widths in the reading of books. This does not mean, however, that one typographical arrangement is read just as easily and as rapidly as any other. Important differences do exist and a genuine improvement in legibility can be made by seeking optimal printing arrangements.

READING REFERENCES

Burtt, H. E., *Psychology of Advertising.* Boston: Houghton Mifflin Company, 1938, Chap. 17, "Typography."

Huey, Edmund B., *Psychology and Pedagogy of Reading.* New York: The Macmillan Company, 1908.

Luckiesh, M., and F. K. Moss, *Reading as a Visual Task.* New York: D. Van Nostrand Company, Inc., 1942.

Paterson, D. G., and M. A. Tinker, *How To Make Type Readable.* New York: Harper & Brothers, 1940.

———, "Readability of Newspaper Headlines Printed in Capitals and in Lower Case." *Journal of Applied Psychology,* 1946, **30,** 161–167.

Pyke, R. L., *Report on the Legibility of Print.* London: H. M. Stationery Office, 1926.

Strang, R., and F. C. Rose, *Problems in the Improvement of Reading in High School and College.* Lancaster, Pa.: Science Press, 1938.

10. ILLUMINATION

DEAN FARNSWORTH

Lt. Commander, USNR; Head of Color Vision Selection and Research, Medical Research Laboratory, US Naval Submarine Base, New London, Connecticut; Color and Lighting Consultant; Chairman of the Technical Committee on Color and Illumination, Illuminating Engineering Society. Formerly: Technical Director of the University Theatre, New York University; Scene Designer and Stage Lighting Consultant.

Opinions or conclusions contained in this article are those of the author. They are not to be construed as necessarily reflecting the views or endorsement of the Navy Department.

Moonlight is adequate illumination by which to find our way along a garden pathway with ease, accuracy, and pleasure, but if we attempt to pick four-leaf clovers the same light is inadequate. Adjustment to the new task has suddenly increased the coordinated activity of billions of cells in the most complex organ of environmental adjustment in the body—the visual system.

So simple a visual act as that of accommodation from a far to a near object involves a forward motion of the lenses of the eyes, an increase in curvature of their anterior surfaces, opposed movements of the three pairs of extrinsic muscles in each eye, and contraction of both pupils. If this accommodation is inadequate to the task because of insufficient illumination there will be inefficiency and fatigue. Continued effort in any improperly lighted visual situation may result in eyestrain, inflammation of the accessory organs, astigmatism, myopia, headaches, indigestion, bodily fatigue, and emotional disturbances. Wittekind (1941) notes more than forty such specific defects as symptoms of eyestrain.

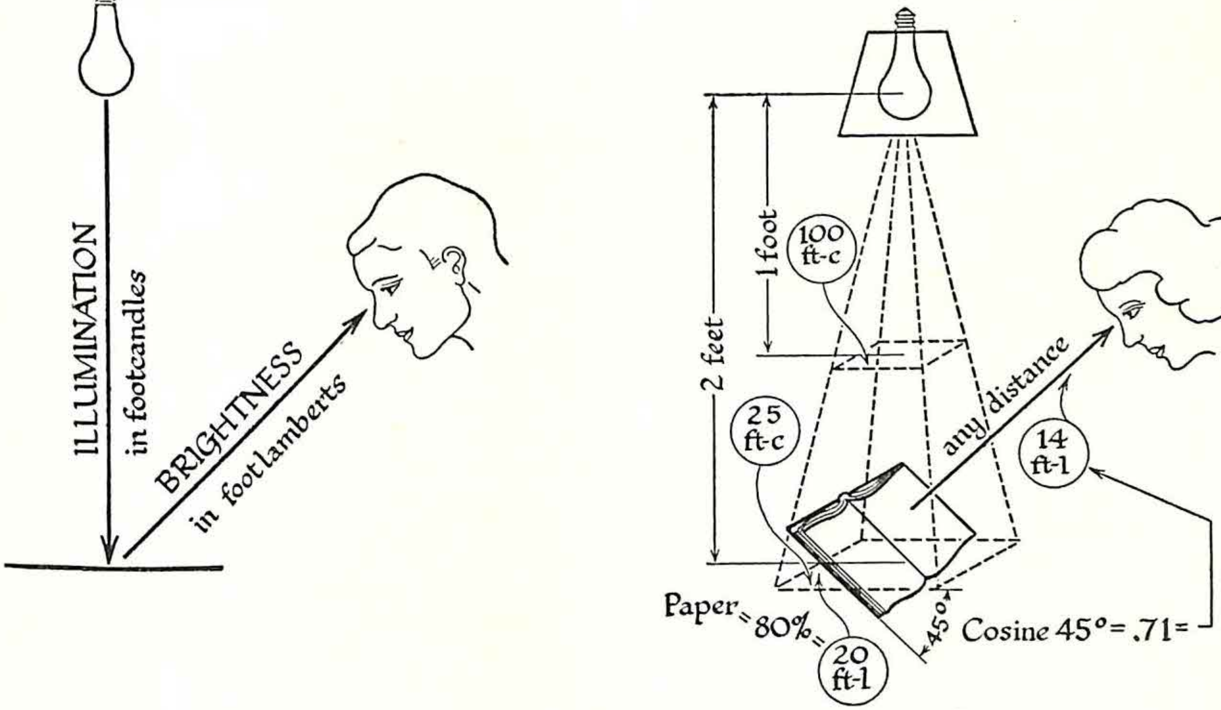
MEASURES OF LIGHT

Water is measured in quarts, light in footcandles and footlamberts. A footcandle (ft-c) is approximately the illumination which falls on a surface held at right angles to a candle flame at a distance of one foot. What splashes back, so to speak, is brightness. Brightness is measured in footlamberts. A footlambert (ft.-L) may be simply envisaged as a brightness equivalent to a perfectly diffusing, perfectly white surface which is illuminated by one footcandle.

As an object is removed from the source of light the illumination decreases as the square of the distance. If the receiving surface is not at a right angle to the direction of the light, the illumination in footcandles is further reduced as the cosine of the angle of deviation. White paper reflects about 80 per cent of the incident light and a black surface from 3 per cent to 6 per cent. The product of (1) reflection for diffuse surfaces in percentage and (2) illumination in footcandles and (3) the cosine of the angle equals the brightness in footlamberts, which is the appropriate measure for visual studies. These measures are shown in the diagram.

VARIABLES OF VISION

The principal variables in any visual situation include: (1) the task, (2) the physiological apparatus, and (3) the conditions of illumination. Certain variables of illumination are: (1) general brightness levels, (2) contrasts of brightness, (3) quality of illumination, (4) injurious light sources, and (5) aesthetics of illumination. These may be related to factors in specific situations: (1) size of the object, (2) the reading situation, (3) color discrimination, (4) contrast of object and background, (5) movement of the object relative to the eye, (6) duration of exposure, (7) state of adaptation of the eye, (8) peripheral vision, (9) aesthetics of room lighting, and (10) the affectivity of color.



A B

Figure 1 (II–10). (*A*) *The Measures of Illumination and Brightness;* (*B*) *the Reduction to Final Brightness Reaching the Eye Owing to Distance, to Percentage of Reflection, and to Angle of Incidence*

Many of these variables are interdependent. For example, in order to resolve a test item under decreased illumination, the size of the item may be enlarged; or the brightness contrast of the item with the background may be raised; or the color may be changed. Likewise, increased size, increased contrast, or increased illumination can be used to compensate for the stimulus loss due to brief exposure.

ILLUMINATION LEVELS

Extreme brightnesses may occur in such ratios as to cause painful glare, retinal fatigue, and general physical strain. Natural illumination may run from less than 1 footcandle to 5,000 footcandles and more, but the upper levels of natural illumination may not be comfortable or safe, nor is it practical to employ high illuminations for indoor work. Many investigations have been devoted to discovering for particular tasks the lower levels of illumination which are consonant with efficiency, comfort, and safety. Adequacy of illumination has been examined by measures of acuity, speed, accuracy, and fatigue. Crouch (1945) has summarized the literature on engineering relationships with acuity, contrast, time, brightness, glare, and defective vision.

ACUITY RELATIONS

Detailed discrimination at the limits of acuity corresponds closely to the brightness acuity curve. Coarse work can be efficiently accomplished in relatively dim light. Adams (1935) concluded that 3 footcandles to 4 footcandles is desirable for tile pressing.

The best illumination for handsetting 6-point type (this size) has been reported as 13, 25, and 20-50 footcandles (Weston and Taylor, 1933; Beutell, 1934; I.E.S., 1936); for needle-threading, 30 footcandles (Troland, 1931); for dental work from 100 to 350 footcandles (I.E.S., 1941); and surgical operations up to 2,500 footcandles (Beutell, 1934; Harrison and Wietz, 1928; Alexander Ickis, 1938).

Reports have been made by Schober (1938) and others that maximum acuity is attainable at levels as low as 10 footcandles but this is contrary to ordinary experience as well as experimental evidence. Lythgoe (1932) demonstrated that such results may be due to failure to control brightness contrast with the surround; with properly lighted surrounds acuity continues to rise beyond 1,000 footcandles.

TIME RELATIONS

The incessantly active eye darts from point to point in saccadic movements of a few thousandths of a second, pausing just long enough to record a significant image. If there is sufficient brightness to satisfy the worker, these pauses last less than ⅕ second; if not, each pause is lengthened or repeated until a satisfactory image is achieved.

Insufficient brightness will result in faulty work and a decidedly slower rate of performance if details of the task are moving in relation to the worker or are vibrating. Clerical checking, assembly of mechanical parts, inspection, sorting, and machine-tending are tasks affected in this manner by too low brightness. The critical level below which errors are noted in sorting mail lies between 8 and 10 footcandles. Both the audience and the participants appreciate from 30 to 100 footcandles on the playing area in game rooms and on the outdoor field, depending upon the speed of the ball. Subway and train lights should exceed 15 footcandles according to the degree of unsteadiness imposed on the riders.

Cobb (1924), Ferree and Rand (1927), and Weston (1945) measured the effect of illumination in terms of speed of performance with similar results. It becomes evident that high efficiency in dealing with fairly coarse detail is possible with low illumination but that the highest efficiencies demand disproportionately higher levels; that when the details of the task are near the visual limen, extremely high levels are required for fair efficiency.

BRIGHTNESS RELATIONS

The use of footcandles of illumination as a measure of brightness for discriminatory purposes is misleading unless the subject materials are very light. Operational conditions should be specified in footlamberts because dark materials present the same visual situation as would be found with lower illumination on high-reflection materials. Operations on men's suits, metal castings, upholstery, and stained woods require from 2 to 20 times as much light on the working surface as would be required for similar work on light-colored materials. Coffee-roasting and soil-analysis require about 5 times the intensity needed for butter-scoring. School blackboards should be lighted at least twice as much as the room.

CONTRAST DISCRIMINATION

When the comparison fields are larger than 2 degrees, acuity is not a factor in contrast discrimination and maximum sensitivity is approached at lower brightnesses than are necessary for size discrimination. Paper and textile chemists are satisfied with local illumination of from one-half to one-tenth of that needed by engravers. Matching could be done at from 1 to 10 footlamberts with fair accuracy with ordinary biparte fields commonly found in laboratory instruments were it not for the dark surrounds of the tube and eye piece.

The various values which have been found for $\Delta I/I$, or just noticeable brightness difference, have been examined by Munsell, Sloan, and Godlove (1933). The fraction probably exceeds one per cent at daylight levels. Values vary for different eyes but the certain fact to be gotten from such studies as Stiles and Crawford (1934) is that the so-called Fechner fraction is not a constant for extended ranges of brightnesses.

CHROMATICITY RELATIONS

The task of color matchers, dye and pigment mixers, color sorters, and graders in all phases of industry and research, depends upon chromaticity discrimination. Discrimination of chroma is nearly uniform in the entire photopic range. It is but slightly higher at 50 footlamberts than at one footlambert. No data are available for hue discrimination, but since value differences are a part of the color-grader's work, and because of the importance of slight differences between samples, there is need for high levels of illumination. Installations for color-graders at the US Bureau of Standards are between 100 and 150 footcandles, according to Nickerson (1939). The Macbeth Daylighting Corporation, which supplies standard light sources for all types of color inspection, use a basic figure of 20 footlamberts from which to compute the necessary footcandles for various working surfaces. The footcandles required would be 20 divided by the average percentage reflectivity of the materials with which the worker is dealing. Experience indicates that to obtain the absolute maximum of chromatic discrimination even higher values are needed.

VISUAL DEFECTS

Acuity, size of pupil, pupillary accommodation, and amplitude of accommodation (focus) decrease with age. Ferree and Rand (1934) increased illumination from 1 and 5 to 25 footcandles and found that increases in acuity and other requirements of good vision were much greater for presbyopes than for normals. The range of distance accommodation was increased for the middle-aged and aged. Consequently they recommend increased illumination for persons with refractive errors, presbyopia, and astigmatism, and suggest the need for light specification according to segregated age groups.

Many color blind individuals complain of glare and visual inefficiency at illuminations which are not disturbing to others. For color blinds who are largely dependent upon rod vision, lighting conditions may need to be carefully controlled so as to bring the maximum tolerable amount of light on their work. According to Abney (1895), the luminosity functions of the color-defective are reduced as compared to the normal, which implies a reduction in their brightness discrimination and acuity functions.

READING

Reading ordinary black print on white paper is principally a task of supraliminal acuity discrimination. Numerous experiments agree in showing a rapid rise in discrimination of just perceptible detail up to 5 footcandles, some gain up to twenty, and less measurable increase thereafter. But it cannot be assumed that proper illumination values for reading can be derived from an acuity curve developed from these data. Ordinary life situations do not demand such exacting acuteness; print is rarely so fine as to be just readable. These facts suggest that less light may be needed than the curve indicates. On the other hand, the curve is a measure of efficiency and not of comfort or visual safety, which suggests that more light may be needed.

From 15 to 25 footcandles have been recommended for general reading in publications of the Illuminating Engineering Society (1938, 1942). The tendency is toward the higher levels. Basing their values on indirect measures of fatigue, Luckiesh and Moss (1937) have recommended from 20 to 50 footcandles and more. Tinker (1939) maintains that these values are unnecessarily high and makes the following summary: "In reading ordinary book type . . . the critical level of illumination intensity, beyond which further increases in brightness produce no further improvement in performance and no further decrease in visual fatigue, is apparently 3 and 4 footcandles. . . . When the visual task is somewhat more exacting, where less than optimum legibility is present, the critical level lies between 8 and 10 footcandles." These differences of opinion can only be resolved by a revaluation of experimental techniques, methods of interpreting data, and the application of suitable criteria of visual efficiency for particular tasks.

FACTORY WORK

All published figures show substantial rise in production when low illuminations are raised. Wyatt (1923) found a 5 per cent loss of output when illumination was reduced to one footcandle for cotton weaving, which is not a very exacting task. Weston (1938) achieved an improvement in output of 5.3 per cent for worsted weavers by doubling the illumination from a maximum of 13.3 to 29 footcandles. He decided that 30 footcandles on fairly dark cloth is necessary for high efficiency. Piston ring production was increased by 25.8 per cent when the illumination was increased from 1.2 footcandles to 14 footcandles. Troland (1931) concludes that the vast majority of industrial operations can be carried out at maximum efficiency with an illumination of approximately 10 footcandles. There is no doubt that prolonged detailed work is aided by ten times this much light. The I.E.S. "American Recommended Practice of Industrial Lighting" (1942) suggests minimum standards for specific applications in industrial interiors.

BRIGHTNESS CONTRASTS

Continual adjustment to changing light levels is highly fatiguing. If, while reading under a single floor lamp, we turn to see what causes a noise in the dark corner, there is a burst of physiological activity: (1) the iris enlarges; (2) probably, as demonstrated in the lower animals, the inner segments of about 12 million cones elongate, about 200 million rods contract, and the pigment granules start to retract into the pigment epithelium; (3) an electrical potential is generated in the eyeball; (4) the retinal reaction becomes more acid; and (5) the cyclic regeneration of the visual purple increases. When the eyes are turned back to reading, the whole process is arrested and reversed. The process constitutes work in a real sense, physical work entirely out of proportion to the size of the organ.

If a task is such that a workman's eyes must repeatedly shift from a dark toolrack to a bright work area, the constantly changing adaptation will produce fatigue, errors, and a slower work rate. The student under his reading lamp, the carpenter at his bench, the shoemaker at his machine, the cashier in the change booth, and the plane pilot at night, if provided with a strong localized lighting will be subject to the fatigue and delay of constant readaptation.

A ratio of 1 : 10 for adjacent brightness is considered permissible by Harrison and Luckiesh (1941) though a committee of the I.E.S. concludes that the ratio should not exceed 1 : 3 for good seeing. In no case should the surround be brighter than the task. It must be noted that the above are brightness, not illumination, ratios; reflectances of background fields may be so low as to require much higher illuminations in order to maintain brightness ratio. Moon (1941, 1942) has stigmatized the harmful contrasts which result from the unnecessary use of dark floors, walls, and furniture. Brainerd and Denning (1941) improved vision in machine tool operation by painting the machines in slightly contrasting colors, e.g., green, buff, yellow, and light grey.

DISTRIBUTION

The principal visual cues by which men and animals perceive their environment depend upon light and shadow. The other cues reside in the special reflectivity of surfaces, called local or object color. A flag is distinguished only by its local color and a bas-relief only by the pattern of shadow.

Diffused light is good for color-matching because it obliterates texture. Reception rooms, drafting tables, theater auditoriums, and picture galleries may achieve effective illumination with indirect fixtures. But directional lighting should predominate where working surfaces are significant, as they are on stairs, motor roads, kitchens, and machine shops.

Neither diffuse nor directional lighting should be used alone to illuminate a large area, for example, a factory room. Diffused illumination from completely indirect fixtures or an overcast sky causes peripheral glare.

GLARE

Discrimination is greatest when the detail is surrounded by a field of approximately the same brightness (Ives, 1921), or slightly lower (Lythgoe, 1932). If certain parts of the visual field are much brighter, glare exists, which constitutes the real lighting hazard in daily life. The "American Recommended Practice of Industrial Lighting," prepared by the Illuminating Engineering Society, classifies glare as "direct" when caused by excessive brightness or brightness-contrast within the visual field, and "reflected" when caused by glossy or specular surfaces in the working plane. Glare may cause fatigue, inefficiency, and risk of accidents. Logan (1941) finds that conditions of ocular comfort may not be most favorable to optimum visibility.

The effect of a bright light near the center of vision is to cause veiling glare which lowers discrimination and obscures detail by dispersing light through the inter-ocular media, inducing blackness over the rest of the field, and raising the general level of adaptation. This is familiar to motorists when the sun reflects from the chromium trim or when they face the afternoon sun or night headlights; to students reading shiny book papers; to repairmen, fishermen, sailors, typists, machinists, and all who are exposed to highly reflecting surfaces. Low-hanging unshaded lamps are among the commonest cause of direct glare.

Remedy lies in the adoption of dull finished reflecting surfaces, removal of the light source from the field of vision, increasing the area and thereby reducing the brightness of the luminaire, or by raising the general level of brightness. "Polaroid" can be used to mitigate reflected glare under certain conditions, especially over water, specular road surfaces, and desk tops. It has many other specialized applications.

Glare is commonly used in the sense of "too much light." Within the normal field of view Harrison and Weitz (1936) recommended that no brightness exceed 1,500 footlamberts for casual viewing; nor 500 footlamberts for prolonged exposure. Such values really depend upon the distribution of light in the visual field. Peripheral brightnesses are irritating, possibly because of the predominance of rods in the retinal areas outside the maculae. Rods have a low threshold and give a glare response at ordinary scotopic brightnesses. Peripheral glare does not interfere with acuity so much as it causes fatigue and discomfort (Luckiesh, 1942). The overuse of indirect lighting produces large, bright areas of ceiling and walls which cause peripheral glare, probably more often than has been suspected. Lights should be kept outside or as far to the edges of the field of vision as possible; large wall areas may need to be reduced in brightness.

Where the elimination of peripheral vision is not hazardous, simple goggle-type shields without lenses give excellent protection.

SPECTRAL QUALITY

IN RELATION TO ACUITY

Because the human lens is not corrected for chromatic refraction, the image from a single point illuminated by mixed or white light is not a single point on the retina. Consequently it has been assumed that acuity would be increased under monochromatic light. Experimental evidence has generally failed to support this belief and the explanation may lie in the absorptive characteristics of the macular pigmentation (Wald, 1945) or in the apparent "yellow-blue blindness" of very small areas of the retina (Hartridge, 1945). Whatever differences may exist between efficiencies under white and monochromatic light are evidently insignificant. For highway lighting no type of light—mercury, sodium, fluorescent, or incandescent—seems to be of special benefit over another (Luckiesh and Moss, 1936). The yellow glasses in so-called "foglights" do not contribute to visibility and red and orange goggles have been shown to possess no "haze-cutting" qualities for the human eye.

Only lights of special mixture, such as purple, composed of red and violet rays, will decrease acuity to a practical extent. Purple spotlights are used for certain stage shows because of the complete inability of the eye to focus on purple-lighted detail. For night lights in bedrooms and sleeping cars, violet provides illumination without emphasizing detail. During blackouts red lights permit maximum discrimination at close range with minimum pickup at distance.

CHROMATIC DISCRIMINATION

The color of an object is dependent upon the spectral quality of the light by which it is viewed. White light or daylight may be anything from the yellow color of sunlight or of a Mazda lamp to the blue of the sky. It is usually considered as something between the two. Certain spectral distributions, designated Illuminant A, B, and C, have been selected, defined, and descriptions given for their reproduction with artificial light by the 1931 International Commission on Illumination. Lamps which approximate the international standards are commercially procurable. It is not sufficient that two lights look alike if they are to be used for color matching and color discriminations; they must have the same spectral composition. Conversely, two textile or pigment colors cannot be said to be matched unless they remain matches under two extremes of color temperature.

No color description has much meaning except in relation to its illuminant. Dress goods which are perfectly matched to accessories under artificial light may not remain matches in the street. A painting executed under north light will suffer serious distortion if the gallery is lighted, as is too customary, with low-wattage incandescent or fluorescents in the strip reflectors. A bluish suit selected by fluorescent illumination may be greenish outside the store. Fresh meats which appear spoiled in a refrigerator lighted by daylight fluorescents may appear excessively red under soft white due to the fact that fluorescents are essentially a mercury lamp with a highly irregular spectral distribution.

Nickerson (1940) investigated illuminants for color discrimination to find (1) the illuminant under which color differences would surely be evident; and (2) how well one illuminant duplicates another. For the detection of fine differences such as are desirable in government grading of hay, produce, and soil, she concludes that the advantage lies with illuminants at color temperatures in the region of 7000° Kelvin or above. It appears that the bluer the light the lower the level of illumination preferred by color graders (Nickerson, 1948). The Textile Color Card Association and the Munsell Color Company set their standards under Illuminant C. Many retail stores provide fixtures with correction filters designed by manufacturers of scientific light control.

INJURIOUS ILLUMINANTS

The visual fatigue which results from inadequate lighting is neurologically and psychologically detrimental and a potential source of accident. In addition, certain radiations may cause local and immediate injury to the organ of sight itself. With the increasing use of specialized radiation apparatus in the home and factory, these hazards to sight need to be better understood and safeguarded. LeGrand H. Hardy (1940) lists common ocular injuries by radiation as follows: (1) direct view of the sun, concentrated incandescent filament, acetylene welding torch, or other brilliant spot sources; these may cause thermal injury to the areas of the retina which pass in minutes or persist for weeks; (2) continued exposure to an expanse of bright light, such as sunlit snow or sand or sheets of molten glass or metal, may result in only temporary blindness but is dangerous to the tissue of the cornea. Erythemal or sunburn rays do not penetrate far into the tissues and the concentration and absorption of large magnitudes of energy, especially the near infrared, in the cornea results in abiotic lesions. There is evidence that the effects may be cumulative, resulting in cataract and other diseases later in life (Lloyd, 1929). (3) An ultraviolet source need not be viewed directly in order to cause conjunctivitis, a painful inflammation of the mucous membrane covering the front part of the eyeball and lining the eyelids. It may last a day or increase in intensity up to two weeks. Mere exposure of this tissue to quartz-mercury lamps, sunlamps, electric power-flashes, and arc lamps may set up the disturbance. The extremely short radiations of an antigermicidal lamp may irritate the conjunctiva in a few seconds and at close range will affect the exposed skin in a few minutes, though the symptoms do not begin until hours after. Less severely, sunlight will affect unconditioned eyelids in the same way.

EYE PROTECTION

Ordinary glass excludes injurious ultraviolet radiation but sunglasses should be designed to reduce heat and visible radiation as well. From 10 per cent to 20 per cent transmission is acceptable in the visible region, from 5 per cent to 10 per cent in the first octave of the infrared. Glasses should be practically opaque to radiation below 300mμ. These important qualities of sunglasses cannot be determined by observation over the counter and should not be purchased unless the glasses are specifically labeled with the above information. It is equally important that the spectacle lens be so broad or well shielded at the side as to reduce peripheral brightness. Highly colored sunglasses prevent sure recognition of signal lights and a neutral tint renders color values more agreeably. It is absurd to wear sunglasses except in bright daylight and dangerous to use them for night driving.

A wide variety of safety goggles has been developed to protect eyes from injurious radiation in industrial operations. The specific application in each case depends upon a sound job analysis.

PLEASANT LIGHTING

Lighting which causes discomfort or fatigue is obviously unpleasant, but a room may be correctly lighted from the point of view of visual hygiene and working efficiency and yet cause unpleasant feelings. The affectivity of illumination involves many less tangible variables in addition to those already mentioned so that laboratory tests scarcely can be expected to yield complete specifications for pleasant lighting.

ACHROMATIC FACTORS

Flatly lighted areas are much less pleasant than are those upon which there is a play of light, e.g., a gradation from top to bottom or a varied panel treatment. Center-of-the-room lights are murderous to charm. Cove lights in the ceiling, windows, or walls introduce pleasant transitions and should be planned so as to differentiate adjacent walls. Local lighting should diffuse over the room and blend gently into the other lights.

Too little contrast, as previously noted, renders an imperfect sense of form. It gives an impression of greater size, of impalpability, as in the palace foyers of cinemas and expensive dress salons, and sometimes of vagueness or of disquietude. Too much contrast is irritating to the eye and to the disposition. It may have attraction and selling value in commerce but it is a distraction from unity of appearance in home and office furnishings.

There must be a sense of liveliness about the lighting to achieve pleasure from it. Liveliness may be analyzed as consisting of the soft highlights which result from broad or well-shaded sources, accidentals or the cross-play of light from several sources, luminous shadows as the result of light wall reflections or low brightness sources, and backlighting which is the accent of light from behind or at the side of objects. Several varied light sources are necessary to accomplish this, though one type of lighting must appear to predominate. All lights must be shielded from the eye, diffused through large lamp shades or shields, or be reflected from the walls and ceilings over a large area. Low illumination in a window recess, wall translucency, or ceiling panel is invaluable for a pleasurable effect. Direct and indirect lights are best placed in different parts of the room and not combined in one unit. The most brightly lighted areas must be where the things to be seen are located, not above or at the side. Misdirected lighting which distracts from the performers, the sculpture, the paintings, is the commonest fault of music studios, auditoriums, and museums.

CHROMATIC FACTORS

Lighting which is designed for people to live under should be flattering to their skin. The hue-summation, therefore, is limited to the range between yellow, orange and red since other colors emphasize symptoms associated with ill-health, e.g., bluish veins, grey blemishes, and the apparently greenish tint of the epidermis. Local high lighting may be of any hue if its intensity does not destroy the general warm effect, but colors should never be used in full saturation. Even good stage lighting employs only the tints of color.

This principle of stage lighting should be extended to wall and textile color. An unpleasant disorientation results from illumination which is in contrast to the principal surface which receives it. It is, therefore, sound practice to pick up the light with materials which are of nearly the same hue as the light, or vice versa. For a room decorated in blue and grey, it would be a mistake to use ordinary incandescents in opal glassware; daylight blue lamps will give a warmer and more harmonious effect. Soft white fluorescent lamps have an excessive output in the red region, which necessitates considerable pink or peach in the decoration of the stores for which they are designed.

Color in light is not distinguishable apart from the color of the surfaces it illuminates, i.e., the color of the paint or of the dress goods depends upon the color of the illuminant. As regards affectivity, therefore, all color impressions can be treated together and can be modified either by changing the paint or the light.

Preferred colors seem to be largely the result of training and associative factors. Guilford (1933) believed he found sufficient agreement to indicate a basic biological cause of likes and dislikes, though culture patterns cause wide differences from one year to another and between the sexes. He reports (1934) a "primitive and fundamental" system in which blue is preferred over yellow and a secondary system of preference for red, green, and blue over yellow, blue-green, and red-violet.

The preferred combinations or color harmonies of academic tradition in art have not been validated by experimental tests. Teachers of art stress the importance of complementary colors, and as a second choice, hues that are very close together on the color circle (analogous colors) rather than those with a medium amount of difference. Guilford (1936) found otherwise and proposed a quantitative law by which the affective value of a combination of two colors could be predicted from the separate affective values of the two components. A study by the author of eighty museum pieces, ranging from oriental rugs to famous paintings showed that their dominant hues were not complementary. There is no evidence that balance on neutral is necessary to harmony. So-called "scientifically measured systems of color harmony" are so much patent medicine. This statement must not be interpreted to include such a system of color standards as the Munsell Company provides for color research. Triads consisting of almost any three unlike colors may elicit either pleasant or disagreeable responses depending upon their hue, value, and chroma distributions. A rule which governs these distributions is not discoverable in the thousands of articles which have been written on color harmony.

In spite of all these negative findings, two recommendations can be made concerning preferred colors: (1) if the illumination deviates substantially from white, more than one color should be used, e.g., not even dance halls are agreeable if lighted entirely with pink, or amber, or any other one color; and (2) there should be a sequential relation between the colors; hues must not be of equal saturation and brightness, e.g., a series might increase from weak yellow general illumination through stronger orange cove lights to bright rose in local areas, diminishing to low violet for alcove lighting.

No reliable evidence supports any of the following popular conceptions: (1) that so-called warm colors will cause a feeling of increased room temperature; (2) that blue objects feel colder than red; (3) that specific colors of lights will remedy pathological disorders; (4) that eye treatments with spectral light will cure diseases of the eye or affect emotional unbalance; or (5) that specific hues will induce certain moods.

Color association and symbolism are apparently culture patterns. Not black, but white, is the color of death and mourning in Japan, and of devils and evil in Africa. A carefully controlled verbalization study by Wells (1910) with American college students showed the peak of excitation in the orange-red region of the spectrum, of repose in the green, and of gravity in the violet. But such mood effects are not so deeply rooted that they cannot be eradicated. A green room lit with cool lights may be made affectively as warm as an amber-lighted maplepanelled library, if sufficient attention is given to light distribution, contrasts, transition, and diffusion.

OPTIMUM CONDITIONS

Hygienic illumination should provide for effective, comfortable, safe, and pleasant vision. For physiological ease of performance, the intensity, quality, and distribution must be suited to the task. When all conditions are optimal, 10 footcandles constitute ample illumination. This level should be modified, usually upward, for liminal tasks and visual defects. Direct glare, reflected glare, and brightness contrasts within the field of vision should be held near the minimum which is consonant with efficiency. The eye should be protected from excessive and injurious radiation. The aesthetics of lighting is influenced by the social patterns of preference which, when satisfied, contribute to effective and comfortable vision.

READING REFERENCES

Crouch, C. L., "The Relation between Illumination and Vision." *Illuminating Engineering,* 1945, **40,** 747–784.

Hardy, LeGrand H., "The Eye as Affected by Illumination." *Transactions of the Illuminating Engineers' Society,* 1934, **29,** 364–384.

Luckiesh, M., and F. K. Moss, *The Science of Seeing.* New York: D. Van Nostrand Company, Inc., 1937.

Lythgoe, R. J., *The Measurement of Visual Acuity.* London: Medical Research Council, 1932.

Tinker, M. A., "Illumination Standards for Effective and Comfortable Vision." *Journal of Consulting Psychology,* 1939, **3,** 11–19.

11. VENTILATION

KINSLEY R. SMITH, PH.D.

Professor of Psychology, Pennsylvania State College. Formerly: Labor Statistician, Philco Radio Corporation; Technical Consultant to the Research Laboratories of the American Society of Heating and Ventilating Engineers; Project Director of Project N-105, and Contractor's Technical Representative of Project N-117, of the National Defense Research Committee.

Maintenance of normal body temperature, a condition necessary to the well-being and comfort of the organism, is dependent upon the balance between heat production and heat loss of the body. Heat is produced in the process of metabolism and is lost by: (1) convection—when heat passes from the body to the air directly; (2) perspiration and evaporation—in which moisture comes from an underlying part of the body to its surface and evaporates, the heat of vaporization being withdrawn from the body; and (3) radiation—the passing of heat, in the form of radiant energy, from the body to cooler surfaces in the environment. The cooling effect in convection varies with the temperature and velocity (rate of movement) of the air. The cooling effect in perspiration and evaporation is dependent upon the temperature, humidity, and velocity of the air. In radiation, the cooling effect varies with the temperatures of environmental surfaces (for example, walls) relative to body temperatures. Respiration is effective in cooling the body, since it involves heat loss through convection and evaporation, as cool air is drawn into the lungs and expelled both warmed and moisture laden (Yagloglou, 1925).

When the temperature of environmental air and objects approaches or exceeds that of the body, and when the air nears complete saturation with moisture, the body can no longer dissipate its heat, or does so with difficulty and at a slow rate. As a result, its own temperature rises and other alterations occur in the human organism. Responses to atmospheric conditions of various degrees of severity may be divided into four not entirely independent categories, as follows: (1) physiological functioning, such as changes in pulse rate and pressure, blood count, distribution of blood, perspiration, respiration, and body temperature; (2) feelings of warmth and comfort; (3) psychological test and work performances; and (4) health and physical well-being.

VENTILATION MEASURES

Various instruments are used to measure the essential atmospheric qualities to which responses classified in the above four ways occur.

DRY-BULB AND WET-BULB THERMOMETERS

Ordinary air temperature is obtained with a mercury thermometer and the reading is referred to as dry-bulb (d.b.) temperature. Humidity is measured with such a thermometer in conjunction with a thermometer which has a moistened wick over its bulb and is known as a wet-bulb thermometer. An air current is passed over the wick of the wet-bulb thermometer, at a rate of 900 feet per minute or faster, causing the reading to drop as heat is absorbed in vaporization. Unless the air is completely saturated with moisture, the wet-bulb reading is lower than that of the dry bulb. The lowest reading obtained is the wet-bulb (w.b.) temperature. With both d.b. and w.b. temperatures determined, relative humidity (relative to saturation at a given d.b. temperature) can be readily obtained from charts or tables (Yagloglou, 1925).

THE KATA-THERMOMETER

An instrument known as the Kata-thermometer (Hill, 1914) is used to determine air velocity, although there are more modern anemometers for the purpose. The Kata is an alcohol thermometer with an enlarged bulb and a stem marked at 100° F and 95° F. It is heated to a temperature above 100° F, and the number of seconds required for it to cool from 100° F to 95° F gives a measure of air velocity when interpreted in terms of the calibration factor of the instrument and its cooling formula (American Society of Heating and Ventilating Engineers, 1940).

When used with a wet-bulb, the Kata-thermometer indicates cooling power as determined by convection, radiation, and evaporation, and provides a single instrument for measuring the cooling power of the environment.

THE GLOBE THERMOMETER

Vernon (1930) devised this instrument which is used to measure environmental radiant heat. It consists of a hollow copper sphere 6 inches in diameter, painted black and enclosing a thermometer. Surfaces in the vicinity of the sphere whose temperatures are higher or lower than the temperature of the atmosphere produce changes in the temperature inside the sphere. Differences between the globe temperature and the surrounding air temperature serve as a basis for an estimation of radiation when allowance is made for convection effects of air velocity (Bedford, 1937).

EFFECTIVE TEMPERATURE

Since atmospheric qualities work in combination to produce their effects, efforts have been made to express variations of such possible combinations in terms of single figures or values. Effective Temperature, which is a combination of d.b. and w.b. temperatures, is such a value (Houghten and Yagloglou, 1923). A d.b. temperature in still air paired with the same w.b. temperature is given an E.T. value corresponding to the readings of both thermometers. Thus 85° F d.b. combined with 85° F w.b. equals E.T. 85°. Other combinations of d.b. and w.b. temperatures which have been experimentally found to produce the same perception of warmth and to have approximately the same effects upon the physiological functioning of the subjects as the particular combination of d.b. and w.b. cited, are also indicated as E.T. 85°. For example, E.T. 85° is obtained with 87° F d.b. and 84° F w.b.

More recent experimentation in the laboratories of the American Society of Heating and Ventilating Engineers has resulted in the production of charts in which the factor of air velocity is included with d.b. and w.b. readings in defining Effective Temperature [see Fig. 1 (II-11)]. Radiant temperatures are not included in Effective Temperature, although Houghten, Gunst, and Suciu (1941) have determined adjustments necessary in E.T. to compensate for introduction of given radiant heat intensities in maintaining a certain feeling of warmth.

PHYSIOLOGICAL FUNCTIONING

Adverse effects of severely warm atmospheric conditions were noted by McConnell and Houghten (1923), who obtained physiological measurements from laboratory subjects under conditions as severe as 112.5° F d.b. and 112.5° F w.b. The subjects were able to endure these temperatures for only 35 minutes and their pulse rates rose to as high as 160 beats per minute. Dill, Edwards, Bauer, and Levenson (1931) reported the breakdown of laboratory subjects usually within an hour of very hard physical labor in still air at 95° F d.b. and 79° F w.b. (E.T. 85°).

The relationships between various lengths of exposure to still air conditions ranging from 77.5° E.T.

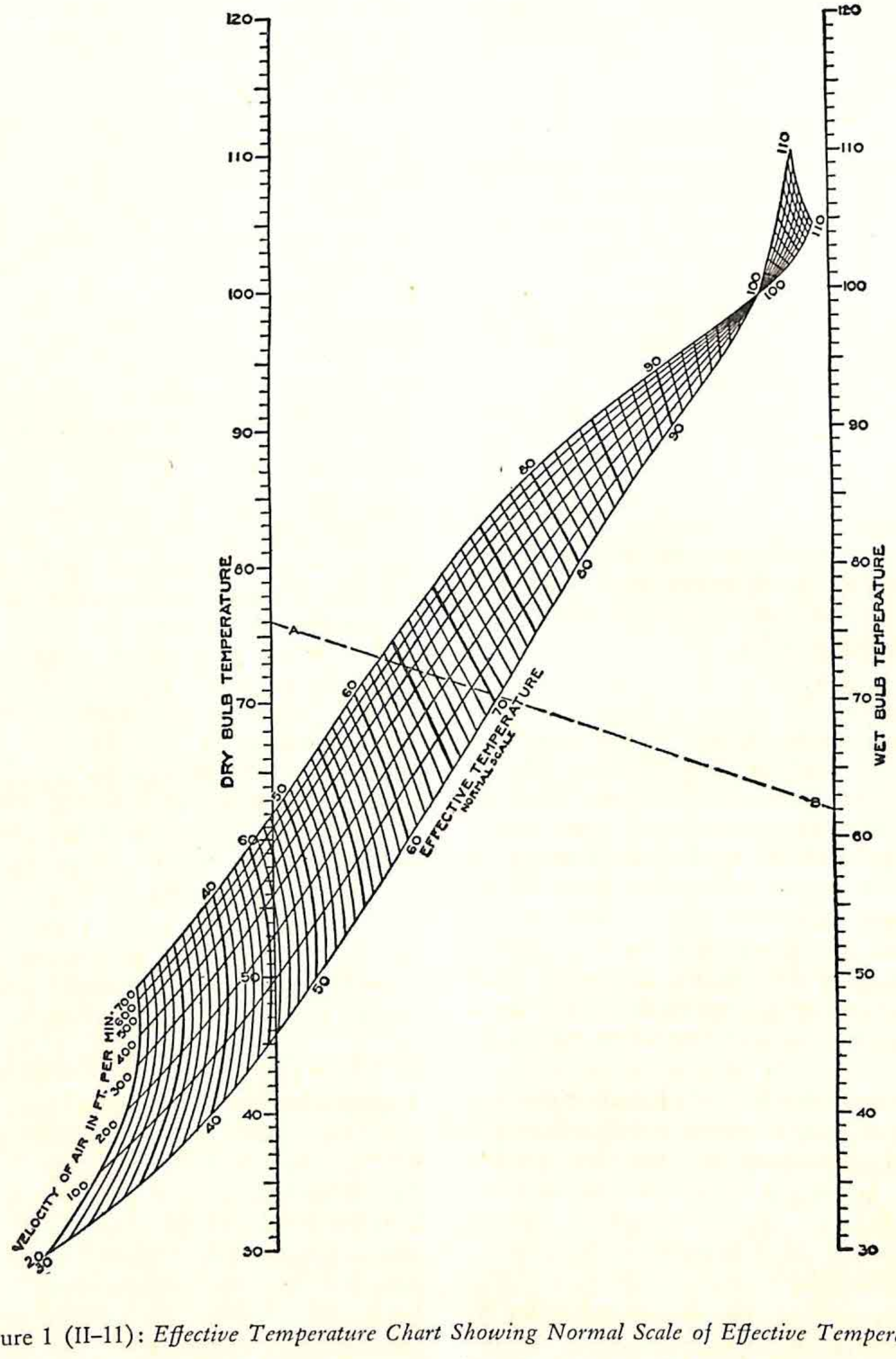


Figure 1 (II–11): *Effective Temperature Chart Showing Normal Scale of Effective Temperature*

Applicable to inhabitants of the United States under following conditions: (*a*) clothing: customary indoor clothing; (*b*) activity: sedentary or light muscular work; (*c*) heating methods: convection type, i.e., warm air, direct steam or hot water radiators, plenum systems. From the American Society of Heating and Ventilating Engineers, *Transactions,* **38,** 1932.

to 92.5° E.T. and physiological reactions in persons performing light work were determined in a study of the American Society of Heating and Ventilating Engineers in 1939 (Fleisher and others). The subjects were engaged for several hours a day over a period of two summer months in an activity requiring them: to stand three fourths of the time, to give rather persistent attention, and to move about a room and move light weights on a table at least once every 2½ minutes. The average pulse rate following 3 hours of work at 77.5° E.T. remained practically unchanged from that of the control rest period at a comfortable temperature. With similar exposure at 92.5°, E.T. pulse rate rose to 50 beats per minute above that of the control period. Body temperature was likewise practically unchanged after 3 hours at 77.5° E.T., whereas at 92.5° E.T. it rose 1.5° F above that of the control period. For the various work periods with E.T. above 77.5°, both pulse rate and body temperature showed an increasing rate of rise for each unit of increase in E.T.

Other studies in the same laboratory, involving heavy physical work under similar conditions, revealed an average gain in pulse rate of more than 50 beats per minute and a rise of 1.5° F in body temperature after 3 hours of exposure at E.T. 84° (1946). These studies were not intended to establish the degree of severity of physiological reactions that can be experienced without adverse effects. However, the investigators offer the recommendation, based upon their observations of the reactions of these subjects, that atmospheric conditions and duration of exposure should never be so severe as to induce a rise in body temperature of more than 0.5° F on the average among physically fit workers acclimated to heat, and never more than 1.5° F among those workers occasionally exposed to heat. This would indicate an E.T. of approximately 80° as the upper temperature limit for persons performing any considerable physical work, and effective temperatures in the vicinity of 70°, or somewhat lower as suitable.

It should be noted, however, that Eichna, Ashe, Bean, and Shelley (1945) have shown that healthy, unclad, fully acclimated military personnel were able to march at a fixed pace carrying twenty-pound packs for 4-hour periods under E.T. 92.5° without appreciable adverse effects. This activity could be carried on only with considerable difficulty at E.T. 94°, and was found impossible at E.T. 96°. A significant finding was that at such extremely high temperatures variations within a relatively narrow range of wet-bulb temperature separated tolerable from intolerable work conditions.

FEELINGS OF WARMTH AND COMFORT

Relations between various atmospheric conditions and comfort and satisfaction are complicated by such factors as the activity of the subjects, the adjustment to prevailing climatic conditions, the attire worn, and the season of the year. Findings must be stated in relation to seasonal and climatic conditions.

During the winter Houghten and Yagloglou (1923) exposed more than one hundred normally dressed male and female Pittsburgh subjects, engaged in card playing and reading, to systematically varied Effective Temperatures and obtained judgments concerning comfort. The majority of subjects rated E.T. 66° (69° d.b. and 60° F w.b. for example) as very comfortable and 97 per cent rated it at least as *comfortable*. In a later Pittsburgh study eight resting male subjects preferred a room temperature of E.T. 67.5° (73° F d.b. and 58° F w.b., for example), and this is regarded as a more desirable winter temperature (Houghten, Gunst, and Suciu, 1941). Investigations concerned with optimal requirements for summer air conditioning have indicated an optimal E.T. of 73° in Pittsburgh and Texas, and 71° in Toronto (Newton and others, 1937).

Subjective reactions of 275 employees in a Minneapolis plant to variations introduced in summer air conditioning established a preference for E.T. 72° (76° F d.b. and 67° F w.b. for instance) in the warmer months and E.T. 70° (72° F d.b. and 67.5° F w.b. for example) during the cooler spring and fall days (Newton and others, 1937).

Differences among temperatures considered comfortable in various situations suggest that the optimal temperature for a given situation should be determined on the basis of experimental variation from temperatures commonly deemed optimal. It is significant in this connection that in one study subjects performing light manual work experienced maximum comfort at approximately 5° to 10° E.T. warmer (that is, at from 77.5° to 82° E.T.) than the temperatures favored by sedentary subjects (Fleisher and others, 1939). The investigators attribute this rather surprising finding to increased convection due to the moving about which was required, to exposure of a larger body area when standing than when sitting, and to less attention, because of the task being performed, to atmospheric conditions ordinarily productive of discomfort.

A representative investigation of the subjective reactions of office workers to summer air conditioning, both while at work and when changing to warmer and cooler conditions upon leaving and coming to work, indicated that about 66 per cent of the employees of a Milwaukee company found the air conditioning effects good, while about 32 per cent judged them unfavorable (Houghten and others, 1938). The majority of workers have voiced satisfaction and appreciation in many work situations in which adverse atmospheric conditions have been improved by air conditioning.

PSYCHOLOGICAL TEST AND WORK PERFORMANCE

Three decades ago Thorndike, McCall, and Chapman (1916) investigated the effects of atmospheric conditions upon mental test performances, including color naming, cancellation, hard opposites, arithmetic addition, and copying material by typewriter. Two temperature conditions, one of 68° F d.b. with 50° per cent relative humidity (E.T. approximately 64.5°) and one of 86° F d.b. with 80 per cent relative humidity (E.T. approximately 82.5°) were combined with various conditions of outside air supply and air movement up to 45 cubic feet per minute. The performance of twenty college students, working for 4-hour periods on 5 consecutive days, was found to be insignificantly related to the atmospheric environment. Performance on some of the tests was better under the warmer conditions, and on others was superior in the cooler environment, but the differences were small.

Viteles and Smith in 1946 reported the relationship between Effective Temperatures, with still air, of 73° (80° F d.b. and 65.5° F w.b.), 80° (88° F d.b. and 73.3° F w.b.), and 87° (98° F d.b. and 81.5° F w.b.), and the psychological test performance of six male subjects, during eighteen daily 4-hour work periods. The tests included mental multiplication, number checking, lathe tracing, typewriter code, discrimeter, and locations and pursuit (Macquarrie). In general agreement with the study cited of Thorndike and others, differences in performance between E.T. 73° and E.T. 80° were insignificant. At E.T. 87°, however, average losses in output occurred ranging from 3 to 14 per cent on the various tests, and were significant on four of the tests. The subjects subsequently engaged in two 4-hour work periods at E.T. 94° (108.5° F d.b. and 90.3° F w.b.) None of the men was able to complete the first 4-hour work period at this severe temperature, owing to weakness, dizziness, greatly accelerated pulse rate, and other extreme physiological reactions. Four of the six subjects, however, were able to carry on throughout the second work period, indicating rapid partial acclimatization, although significant losses in their output resulted, ranging from 14 per cent to 45 per cent on the various tests. Weiner and Hutchinson recently found that six well acclimated subjects suffered an average loss of 14 per cent in output on a light motor coordination test at E.T. 91° (95° F d.b. and 90° F w.b.), as compared with their performance at "ordinary" room temperature, 65° F to 68° F d.b. (Weiner and Hutchinson, 1945).

The findings of these studies indicate that in the case of partially acclimated subjects during 4-hour work periods, output on psychological tests, including light psychomotor tests, is not altered by variations in Effective Temperature within the range of approximately E.T. 64° to 83°. However, such performance suffers significant loss at a temperature of E.T. 87°, and deteriorates seriously by the time E.T. 94° is reached. More complete acclimatation, achieved through increased duration of exposure at work, might well be expected to reduce somewhat the losses in performance at the higher temperatures.

Some forms of work, such as those in metal industries are made tolerable by forcing cool air blasts over the workers and by frequent rest periods. There are numerous reports of increases in industrial output with cooled atmospheric conditions, but few are supported by quantitative data. Wyatt, Fraser, and Stock (1936) reported slight increases in weaving shed production, namely 1.6 per cent for the morning period and 2.9 per cent for the afternoon period, with the introduction of fans which increased air movement from 46 feet per minute to 147 feet per minute, and dry-kata cooling power from 9.7 to 12.4. However, output of five female electric relay assemblers observed in a Chicago plant showed slight or no fluctuation accompanying seasonal atmospheric changes (Whitehead, 1938). Variation in summer room temperature between 67° F and 90° F, considered singly or in combination with relative humidity ranging from 25 to 70 per cent, was unrelated to output (fans were available and were probably used on the warmer days). Winter room temperature variation between 65° F and 77° F, with relative humidity between 20 per cent and 60 per cent was similarly without effect on work performance.

A relationship between industrial accidents and d.b. temperature seems to have been established by Osborn and Vernon (1922). Male and female workers in two munitions plants were found to have experienced fewer cuts in the d.b. temperature range of from 65° F to 69° F. The accidents rate rose slightly in the 70° F to 74° F range, and sharply above 74° F. Accident frequency likewise increased at lower temperatures, rising steeply below 54° F. Vernon, Bedford, and Warner (1927) added supporting evidence for such a relationship in noting an accident severity rate of 4.9 days lost per one thousand working hours for miners in an environment having a wet-kata cooling power of 10.2, and a rate of 3.2 for those in an atmosphere with 15.2 cooling power.

HEALTH

The beneficial effects of air conditioning upon the health of workers have been described in a number of instances. Viteles (1934) cited a decrease in one company of more than 30 per cent in time lost because of respiratory disorders, such as colds and grippe, among office workers in a year following the installation of air conditioning. Comparison, in the same company, of similar absences in an air-conditioned office and one not air conditioned, revealed more than 50 per cent less time lost in the former. Vernon, Bedford, and Warner (1928) have found increased sickness among miners exposed to d.b. temperatures of 80° F and higher. Factors other than atmospheric conditions admittedly may have influenced these findings which are accepted only as gross evidence of the benefit to health of air conditioning.

The drastic effects of overexposure to extremely high temperatures, such as those above 200° F in certain steel mill occupations, are seen in such familiar diseases as heat cramps, heat exhaustion, and heat retention (Heilman and Montgomery, 1936). Heat cramps, in which the victim may suddenly fall writhing in agony, are characterized, in addition, by cold and clammy skin, flushed face, and dilated pupils. The disease is believed to be due to extreme loss of body salt and dehydration, and is effectively treated by intravenous injection of saline solution. Preventive measures include taking salt tablets by mouth or addition of about 6 grams of salt per gallon of drinking water.

Heat exhaustion involves faintness or stupor in the absence of cramps, with the victim in a condition of extreme shock. Depleted blood sugar and possibly toxic influences are thought to be responsible for this disorder, the incidence of which falls greatly with increase of glucose intake.

Heat retention involves a rise in body temperature by as much as 10 degrees, dry and "burning" skin, livid face, and excitement or mania. The onset follows cessation of perspiration and may be sudden. Treatment includes cooling the patient by such means as bathing and spraying with cool water, and subjecting him to air currents to increase evaporation. It has been suggested by C. A. Mills (1941) that consumption by human beings of food high in thiamin content may possibly help to offset the deleterious influences of great environmental heat.

ATMOSPHERE AND THE INDIVIDUAL

Indoor summer and winter comfort zones for the average residents of this country at rest or sedentary work have been determined (American Society of Heating and Ventilating Engineers, 1946). The average winter comfort zone centers about 66° to 67.5° E.T., the range of temperature judged very comfortable by the majority of individuals tested. It extends down to 63° E.T. and up to 71° E.T., the limits beyond which 50 per cent of persons feel too cool or too warm. The average summer comfort zone centers about 71° E.T. It is bounded by 66° E.T. and 75° E.T., the temperatures beyond which 50 per cent of persons feel too cool or too warm. Since no temperature which may be established can be optimal for all persons, it will always be necessary for some members of a group to make suitable adjustments in their clothing if they are to be comfortable.

It is likely that temperatures somewhat below the centers of these zones are favorable to individual efficiency and comfort in mental and physical work over long periods, particularly as the difficulty of the work increases. However, from the standpoint of short-time emergency performances, it is significant that mental performance is maintained unimpaired at temperatures considerably above those at which comfort declines.

The usual worker may expect his output in medium or heavy physical work to suffer a loss under temperatures ranging from approximately 80° to 90° E.T. Loss will be experienced even in light physical and mental work as the upper portion of this range is approached, and medium and heavy work must be drastically reduced in rate or discontinued. Instead of attempting to maintain high output by increased effort at these adverse temperatures, the worker should rest frequently and in cooler surroundings. It is important that this be done not only in the interest of comfort and general physical well-being but in order to guard against increased accident susceptibility at high temperatures.

READING REFERENCES

American Society of Heating and Ventilating Engineers, *Heating, Ventilating, Air Conditioning Guide, 1946.* New York: 1946, Chap. 12, "Physiological Principles."

Fleisher, W. L.; A. E. Stacey, Jr.; F. C. Houghten; and M. B. Ferderber, "Air Conditioning in Industry." *American Society of Heating and Ventilating Engineers Journal Section, Heating, Piping, and Air Conditioning,* 1939, **11**, 191–194; 225–268.

Heilman, M. W., and E. S. Montgomery, "Heat Diseases: A Clinical and Laboratory Study." *Journal of Industrial Hygiene and Toxicology,* 1936, **18**, 651–667.

Johnson, R. E., "Applied Physiology." *Annual Review of Physiology,* 1946, **8**, 535–558.

Viteles, H. S., and K. R. Smith, "An Experimental Investigation of the Effects of Change in Atmospheric Conditions and Noise upon Performance." *American Society of Heating and Ventilating Engineers Journal Section, Heating, Piping, and Air Conditioning,* 1946, **18**, 107–112.

12. FATIGUE AND ITS ALLEVIATION

RAYMOND A. KATZELL, PH.D.

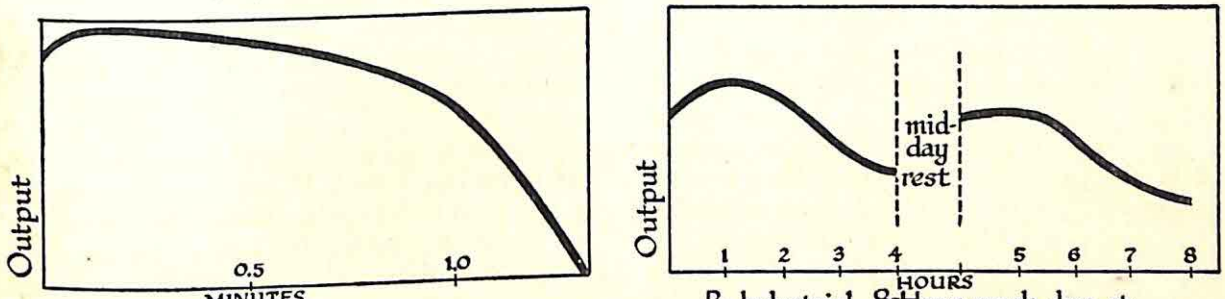
Associate Professor of Industrial Psychology, Syracuse University; Director, Richardson, Bellows, Henry and Company, Inc. Formerly: Associate Professor of Psychology, University of Tennessee; Research Psychologist, Personnel Research Section, AGO, Department of the Army; Instructor in Psychology, New York University and City College of New York; Research Assistant in Psychology, NRC Project on Selection and Training of Pilots, New York University.

Fatigue is one of the chief despoilers of human happiness and efficiency. It is experienced as an unpleasant, strained feeling which interferes with the enjoyment of life; it is expressed as a temporary reduction in productivity that impairs effectiveness and success. Severe fatigue involves physiological and personality changes that may contribute to pathological developments. Increased liability to injury may also be listed among its undesirable characteristics. Although fatigue is an inevitable consequence of prolonged work, research has fortunately disclosed a number of procedures which can help relieve us of its detriments.

NATURE OF FATIGUE

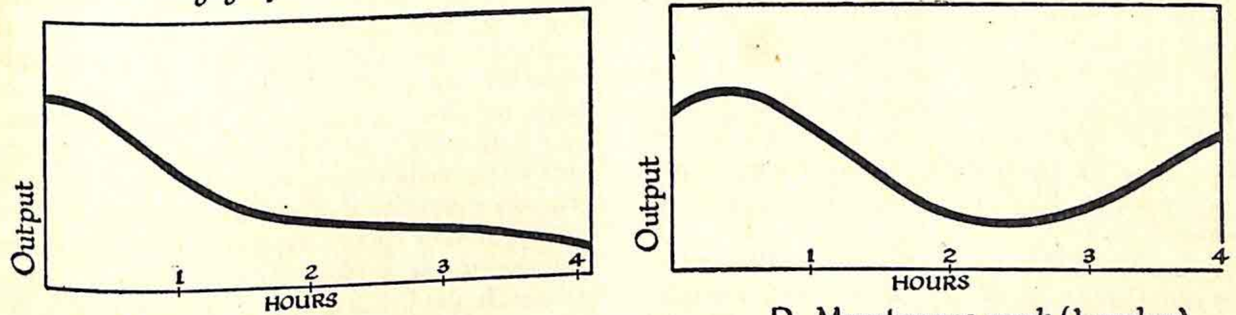
The study and control of fatigue is made difficult by the fact that it is a complex psychophysiological condition having at least three aspects, namely: (*a*) decline in quantity and quality of production for a given amount of effort, the work decrement; (*b*) occurrence of feelings of strain, irritation, and unpleasantness, the feeling decrement; (*c*) a disturbed physiological state characterized by an acceleration of catabolic or energy-releasing processes and the relative abundance of toxins and other chemical products of these processes. Although all three kinds of changes are found to result from prolonged activity, they do not parallel one another closely in the time or intensity of their occurrence. Nevertheless, a general definition of fatigue can be made which involves all of them. It is that fatigue consists of those consequences of prolonged work which decrease the individual's capacity for further work.

Analysis and control of fatigue are further complicated by the fact that productivity and feelings are influenced by the individual's motivation in such a way that it is often difficult to determine when a particular effect is attributable to a change in fatigue and when it is attributable to a change in motivation. This is particularly true in studies performed in industrial or other actual work situations where careful experimental controls are difficult to impose. For example, it is known that the introduction of rest pauses in the industrial working day produces improvements in output. But it is not known to what extent this can be attributed to alleviation of fatigue and to what extent it results from increased effort on the part of the workers as a consequence of improved morale and greater willingness to work. In short, productivity is dependent, not only on working capacity, which is influenced by fatigue, but also on motivation and effort. Changes in fatigue and in motivation may either reenforce or counteract one another in their effects on performance. The interpretation of the influence of various factors on production requires that their effects on fatigue, motivation, skill, and other variables be taken into account.



A. Ergographic work

B. Industrial 8-Hour work day at complex motor work



C. Mental work

D. Monotonous work (boredom)

Figure 2 (II–12): *Curves of Different Types of Work*

These are smoothed curves representing average trends. The curve for a single person would be irregular.

OCCURRENCE AND MEASUREMENT

The occurrence and measurement of fatigue will be described in terms of the three previously mentioned aspects.

WORK DECREMENT

The decline in production occurring after a prolonged work period is demonstrated by comparing the levels of performance (measured in terms of quality of work, quantity of work, or time per unit of work) existing at various times throughout the period. Such a graph of level of performance plotted against time spent at work is called a work curve. Representative work curves are shown in Fig. 2 (II-12).

The usual work curve shows that production improves during the early stages of work. This is called the warm-up effect. Production then reaches a peak and later declines. In individual curves, occasional irregularities or spurts are usually present, and are to some extent the results of variations in motivation. The length of time that must be spent at work before the work decrement appears depends on a number of factors—the nature, difficulty, pace, and method of work, the worker's fatigability, his motivation, and others.

The procedures to be used in relieving fatigue depend in part on the characteristics of the work curve, so that its determination is frequently important for practical as well as theoretical reasons. The level of performance in the work curve can be measured in terms of the work itself when the work is homogeneous and successive units of it are equivalent in nature and difficulty. But many kinds of work, for example, office work and truck-driving, are not of this sort; the work changes, causing the successive steps to be incomparable with each other. The work decrements in such tasks must be studied by interposing at regular intervals through the work period some kind of task as a test, such as to judge the size of objects, and then plotting the work curve in terms of level of performance on this task rather than in terms of the regular work. This procedure rests on the assumption that the fatigue incurred in the original work is proportionately transferred to the test task. Bills (1943) points out that the extent to which fatigue is transferred from one task to another depends on their similarity with respect to (a) neuromuscular processes, (b) directing and controlling attitudes (such as those toward speed, accuracy, and rhythm) and (c) the setting in which the tasks are performed when this is an element producing fatigue in the first task. These factors should be taken into account in the selection of the tasks set as tests. Among the many test-tasks used in the past have been memorizing, reaction time, tapping, tests of muscular strength and endurance, tests of motor coordination and accuracy, and measures of sensory acuity. A survey of numerous studies does not yield consistent results on the value of these tests, perhaps because the test-tasks were not always sufficiently similar in all essential elements to the work itself. The problem of fatigue tests should be studied further because of their importance in determining curves for heterogeneous work. The experiments of the US Public Health Service on interstate truck drivers (Jones, Flinn, Hammond, et al., 1941) are illustrative of this approach.

Decrement in Motor Work. All work involves the functioning of sensory, neural, and muscular units of the body. However, the term "mental work" has been applied to performances requiring relatively little muscular activity, and the term "motor work" to those involving considerable amounts of muscular activity. The distinction, although rather arbitrary, is convenient. Frequent use is made of these terms, because fatigue characteristics vary to some extent with the amount of muscular involvement.

The curve for motor work was first studied scientifically more than half a century ago by use of the ergograph, an instrument that records the output of a group of muscles in some part of the body acting against the force of a spring or a weight. The decline in output is shown in the reduction of the force of the movements.

Ergographic work, of course, involves fewer muscles and stimuli, and also simpler coordinations than does motor work ordinarily done in real-life situations such as in industry and athletics. However, ergographic work curves have the general characteristics of work curves as described above. Ergographic studies thus provide a number of clues to the general conditions promoting or postponing fatigue in motor work; some of the pertinent experimental results follow. Fatigue, shown by the output decrement, sets in more quickly when the work is more difficult (greater force required), and when the rate of exercise is more rapid. A greater total amount of work can be done before complete fatigue occurs when the work involves a moderate load than when it involves a heavy load. The sooner a decline in output appears in a given ergographic task, the less marked is the rate of decline after its appearance, suggesting the influence that the conservation of muscular strength has on the prolongation of productive work. Prolonged work on the part of one group of muscles reduces the capacity for work of other muscles, although to a lesser degree than that of the muscle group which was actively at work. All these facts point to the significance of excessive intensity, load, and pace of work in inducing fatigue in motor performances.

Another interesting disclosure of ergographic research is that the form of the curve of motor work differs for different individuals, although each person's work curve is fairly constant, even for different parts of the body (Yochelson, 1932). This fact suggests the possibility that workers should be placed on jobs having energy requirements that are synchronous to the workers' own work-curve characteristics. It also indicates that an individual, in order to increase his own personal efficiency, might do well to determine his typical curve of work and schedule his activities and rests accordingly.

Decrements in output also occur in light motor work emphasizing speed rather than strength of movement, such as typewriting or tapping. Here the decline usually appears soon after the task is begun, but a level of output is soon reached which can be maintained with little further decline for a relatively long time.

Decrements are found, not only in quantity and speed of motor work, but also in accuracy. This has been demonstrated in typesetting (Muscio, 1920), handwriting (Robinson and Bills, 1926), and numerous other tasks.

Fatigue in muscular work in industry has likewise been studied in terms of decrements in quantity, speed, and accuracy. Output decrements in industrial jobs which are motor in nature and nonmonotonous commonly begin about the middle of the morning (or first half of the work day); the midday rest eliminates some of this decrement, but it reappears toward the middle of the afternoon (or second half of the work day). The peak of output is usually lower in the afternoon than in the morning, and the decrement is commonly more pronounced toward the end of the day. A typical industrial output curve is shown in Figure 2 on page 75. Individual difference in fatigability also have been found in industrial work.

The conditions of motivation do not remain as constantly high in industrial jobs as they do in short laboratory experiments on fatigue. This is particularly true where the work is monotonous or otherwise not satisfying to the worker. Work curves under these conditions frequently show characteristics quite different from those in which fatigue is the sole factor inducing the decrement. This point will be discussed later in this article under boredom.

DECREMENT IN MENTAL WORK. Decrement in mental work occurs in quantity, speed, and accuracy of output. The time at which the decrement occurs and its extent depend on the nature and duration of the work, and the characteristics of the worker. These facts have been disclosed through research on various kinds of mental work, including addition, multiplication, reading, memorizing, sentence completion, judging compositions, and reasoning. Part of the decrement in amount of work done in intense mental tasks seems to result from occasional brief intervals of about a second or two in length of time, during which the worker does not produce (Bills, 1931). These intervals, which are known as blocks, become longer and more frequent as the time spent at work progresses.

Amount and rate of decline of output in intense mental work are surprisingly slight when the worker is well motivated and in good condition. For example, Arai (1912) was able to carry on rapid mental multiplication of 4-place numbers by 4-place numbers continuously for about 12 hours per day on 4 successive days. The rate of output toward the end of each daily session was still about half of what it had been soon after the beginning of the session. A recent repetition of this experiment, with more individuals participating, confirmed this general finding (Huxtable, White and McCartor, 1946). Poffenberger (1928) found that over 5 hours of intense work at addition problems produced an ultimate reduction in output of 20 per cent, and that the same amount of time spent judging compositions resulted in practically no loss in output. These and similar results suggest that a considerable portion, although not all, of the seeming fatigue often encountered in protracted study or other everyday mental tasks may be attributed to decreased satisfaction and flagging interest, and that methods for increasing production in such work should be sought partly in methods for maintaining the motivation of the worker.

PHYSIOLOGICAL ASPECTS

Muscular exercise is accompanied by a number of widespread physiological changes. Among them are increased pulse rate, increased blood pressure and pulse pressure, changes in the differential blood-cell count, increased sugar content of the blood and urine, increase in the lactic-acid content of the blood, increased lung ventilation and increased metabolic rate. The general picture is one involving a more rapid utilization of energy and the greater prevalence of chemical by-products of the energy-releasing physiological processes. The physiological changes are less pronounced in lesser degrees of muscular activity, and are rather slight in mental work.

It should be recognized that physiological changes take place practically as soon as work begins, and are not particularly correlated in time or amount with the work decrement. There is, therefore, no clear-cut physiological index of fatigue, in the sense of a condition which appears or becomes aggravated only after work has been carried on for a prolonged period. Certain indications of this sort have been tentatively identified, but not unequivocally established. These include marked augmentation of blood pressure and pulse pressure (Moore, 1942), submaximal relative carbon dioxide content of the blood at a time when the muscles are contracting with maximum strength (Polakov, 1925), and higher rates of oxygen consumption and carbon dioxide exhalation than those occurring in the same type of work when performed while the worker is still fresh (Andrieu and Fréchinos, 1938).

Even though no physiological criterion of fatigue has been clearly established, measurements of physiological changes do reflect the strain to the worker. They therefore have value in studies designed to identify the work methods which are least taxing to the worker (*e.g.,* Crowden, 1928).

FEELINGS OF FATIGUE

Protracted work results in a feeling decrement as well as a work decrement. As work progresses, one becomes more and more aware of sensations of strain, tension, and even pain, as well as a growing feeling of unpleasantness and irritation. These subjective aspects of fatigue are usually studied by having the worker rate his feelings. There is a tendency for the feeling decrement to develop earlier and with more intensity than the decrement in output (Poffenberger, 1928). However, there are times such as in highly absorbing tasks like athletic contests, when the performer may be scarcely aware of these feelings of fatigue even after his output has begun to deteriorate.

BOREDOM

Boredom, which likewise has a prominent component of feeling, is sometimes confused with fatigue. However, the two conditions can be differentiated with respect to both their subjective and their productive characteristics. Boredom is represented subjectively by a lack of interest in the work being done, by strain, depression, and a desire for a change in activity rather than for rest as in the case of fatigue. The individual work curve in boredom is typically markedly irregular, since it fluctuates with the waxing and waning of interest. It often exhibits a maximum decline in output toward the middle of the work period, and a later return almost to the initial level of production (see Fig. 2, on page 75). The latter phenomenon is probably due to the realization of approaching relief, and perhaps in certain instances to the desire to make up for lost time. The output decrement in boring work is not necessarily transferred to other tasks involving the same muscles (Karsten, 1928), as in the case of fatigue. In light, boring work, changes in output are closely related to changes in the feeling of boredom (Barmack, 1937; Wyatt and Fraser, 1929), considerably more so than the usual relation between the work and feeling decrements in fatigue.

Many repetitive simple jobs require restriction of the range of attention. Poffenberger (1942) believes this to be an important source of boredom. In such situations the worker may try to introduce diversity in the task, and in this way may succeed in reducing effects of boredom. Repetitive work, however, may be relatively free from boredom if the worker recognizes that it is leading him to an objective. Likewise, nonrepetitive activity may be boring if it does not appear to be worth while to the individual, for example, when a movie is neither entertaining nor educational to the viewer. Boredom, then, must be considered, not only as a reaction to uniformity or repetitiveness, but also as a reaction to situations which are ungratifying to the individual's wants and motives.

The relief of boredom can be achieved by stemming its two main sources. Repetitiveness and uniformity can be relieved by variation of work methods, rotation of duties, broadening the scope of a job, playing music, introducing rest periods, or any other way of diversifying the task. Of course, reason should be used to insure that these diversifying procedures will not actually interfere with efficient performance of the job, or else the cure may be worse than the disease. The ungratifying nature of a task can be counteracted by better clarifying its meaning and objectives, introducing frequent incentives, rendering frequent reports of progress, and by other procedures designed to satisfy more adequately the worker's wants through his performance of the work.

As in everything else, people differ in their susceptibility to boredom in a particular job. Persons of high intelligence and extroversion are somewhat more prone to boredom in simple repetitive work than are those who possess these characteristics in less degree (Wyatt and Langdon, 1937). Factors of this kind should be considered together with the person's interests in the assignment of jobs to workers, if boredom is to be held to a minimum.

In many jobs, output and feelings probably reflect both fatigue and boredom, which complicates their understanding and control.

ALLEVIATION OF FATIGUE

Fatigue can be reduced considerably by taking steps to minimize its occurrence and by providing for its dissolution after it appears. Procedures for accomplishing these objectives are discussed below. It should be recognized that most of these procedures improve performance not only because of effects on fatigue but also because of their influences on boredom, motivation, skill, and other aspects of behavior.

REST

Rest is the basic factor in recovery from fatigue and is also important in the resistance to further fatigue. It is essential that the rest following fatiguing work be sufficient to permit practically complete recovery, otherwise the residual fatigue will accumulate and necessitate relatively longer rests to recover from subsequent work sessions (Manzer, 1927). The tendency for daily industrial production to decline during the latter part of the week (Polakov, 1921) may be due in part to undissipated fatigue remaining from the earlier days of the week.

The time required for recovery from muscular fatigue cannot be categorically stated, for the requirement is greater for larger muscles, for work done while the muscles are already partially fatigued, and for greater degrees of fatigue. Other factors, like the activity undertaken during rest, and the worker's attitude toward resumption of work also seem to play a part. The evidence indicates, however, that the rate of recovery is more rapid during the early stages of rest than it is later. Manzer (1927) studied recovery from ergographic work carried to complete decrement and found that recovery was about four-fifths complete after 5 minutes' rest, but that fatigue remained to some extent even after 20 minutes of rest. The recovery from the physiological acceleration accompanying motor work is also gradual (Schubert, 1932).

Recovery from mental work was studied by Kirby (1942). The task consisted of solving rational problems for 30 minutes. It was found that recovery of production ability was 87 per cent complete after a rest of 7 minutes. Likewise, there was 80 per cent recovery from the blocking effects and about 72 per cent recovery from the feeling decrement after 7 minutes of rest. Longer rests produced slumps in production ability, resistance to blockings, and feeling. These slumps probably were due to changes in motivational factors, such as a decline in readiness or desire to resume work. There was a gradual recovery from the general physiological acceleration induced by work; recovery in this regard was more than 40 per cent complete after 7 minutes of rest, and 85 per cent complete after 20 minutes. Kirby's findings are generally consistent with Fryer's (1940) theory that most of the recovery in performance from fatigue resulting from short periods of mental work takes place during the first 10 minutes of rest, and that the remaining recovery proceeds at a slower rate.

Sleep. Sleep is the most complete form of rest attainable, and is of basic importance in recovery from fatigue induced by daily work and activity. Laboratory studies have shown that going without sleep for as long as 100 hours may result in only slight losses in effectiveness. The test tasks in these studies were generally short, so that it is possible that the results reflect certain motivational compensations by the individual. This possibility is supported by the study of Laird and Wheeler (1926), who found that loss of 2 hours' sleep per night produced no loss in output in mental work, but that each unit of work required about three times as much energy as that required after normal sleep. In complex tasks requiring considerable alertness, performance has been shown to decline significantly during the course of 50 hours of sleeplessness (Clark et al., 1946). On the whole, broad experience suggests that sleep is essential for life and health, and that the person who persistently has too little of it will eventually lose efficiency.

The amount of time spent in daily sleep varies, probably depending on habit as well as actual need. The average period of sleep per day for adults is between 7 and 8 hours, and the vast majority get somewhere between 6 and 9 hours. Probably only the exceptional person can get along at top efficiency with less than 6 hours of sleep per day.

Much involuntary insomnia results from poor habits associated with going to sleep. Regularity of time, procedures, and environmental conditions concerned in going to bed are fundamental in inducing prompt sleep. Relaxation of the muscles is also an important aid in this connection, as is the avoidance of intense thinking just prior to bedtime.

Rest Pauses. The interjection of occasional short rest pauses during extended work periods has been found in numerous instances to improve total production by from 5 to 20 per cent. The beneficial effect of rest pauses stems from a combination of five sources. (*a*) Rest pauses provide opportunities for partial recovery from fatigue. (*b*) They help reduce energy expenditure in work (Simonson and Enzer, 1941). (*c*) They may reduce the time wasted in unauthorized rest. (*d*) They break up uniformity, thus relieving boredom. (*e*) They increase morale and motivation of workers (Roethlisberger and Dickson, 1939).

The distribution of rest pauses with respect to their duration, location, and frequency appears to influence their effectiveness. It is impossible to generalize fully on these matters, since their influence depends on the kind and duration of work and the nature of the work curve. However, it seems that the duration of the rest pauses should be in proportion to the length of the work period. Graf (1922), for example, found that in a mental task the optimum duration of the rest pause was 2 minutes for 40 minutes of work and 5 minutes for 80 minutes of work.

Where the work curve is approximately the same as that shown for industrial work in the figure on page 75, the optimum location of a single rest pause in each work period is generally found to be just prior to the decline in output. This location has the dual advantage of not interfering with the warm-up process and of prolonging the period of maximum production, since this usually immediately precedes the decline.

Bills (1935) states that the number and duration of rest pauses should be inversely proportional to each other. Probably rest pauses should be introduced relatively frequently when the work is highly strenuous or boring. In such instances, care should be taken to see that the pauses do not interrupt integral sequences of work.

What is done during the rest pauses is also important. Wyatt (1927) reported that a 15-minute rest during 2 hours of mental work was more effective when it consisted of relaxation than when the pause was unrestricted or when it was devoted to mild physical exercise or mental recreation. Experience in industry has also suggested that relaxation is the most profitable way to spend the rest pause (Mayo, 1924; Hersey, 1925). However, where the work is light and boring in nature, rest pauses should provide for changes in scene, conversation, and other diversifying activities.

Caution should be observed in preventing the disruption of the set toward work by the interposition of rest periods, or else the worker may be less productive when he resumes work than he was before he stopped (Graf, 1922, 1928; Kirby, 1942). Having the rest pauses at regular times and of definite length should help offset this unreadiness to resume work.

LENGTH OF THE WORK PERIOD

Fatigue in its very nature is a function of the time spent at work; as the work period increases in length, fatigue becomes more pronounced. Fatigue may be reduced and total production increased by shortening excessively long periods of work. The pioneer studies along these lines were performed in British munitions factories during World War I. Numerous instances were found where increases in total production followed moderate reduction in the number of hours worked per week. In one department, for example, a reduction in the weekly working hours from 58 to 51 was followed by a production increase of 22 per cent. Similar experiences were reported during World War II. After the fall of France in 1940, the great need for industrial production in Great Britain resulted in marked increases in working hours there. There was a prompt increase in production, followed in the next few months by a general decline (Industrial Health Research Board, 1942). It was stated that the chief cause of this decrease was the additional fatigue and strain induced by the long hours of work. Increases in sickness, accidents, and unauthorized absences contributed considerably to the drop in output. The workers began to feel bored, stale, and irritable. All these deleterious effects were specially marked when weekly working hours exceeded 65 for men and 60 for women. Work-weeks of between 56 and 60 hours were therefore generally adopted in Britain during the War (Industrial Relations Section, 1942). A survey of American industry at about the same time (Brown and Baker, 1942) indicated that in this country the optimum work week for men under wartime conditions was considered to be 48 hours. The differences between British and American results may have been due to differences in prevalent production methods.

The findings discussed above relate to the influence of hours of work on total production. However, a fundamental issue from the standpoint of efficiency is the relationship between hours of work and output per hour. The studies of total production indicate, of course, that when hours of work are excessive, the output per hour decreases. Further light is thrown on this matter by a study of the hourly output during work weeks of 36, 40, 44, and 48 hours (Miles and Angles, 1925). This study showed that although the 48-hour week gave the greatest total production of the four, it resulted in the lowest output per hour. The 40-hour week was most efficient from the standpoint of hourly output, while the 36- and 44-hour weeks had hourly outputs between those of the 40- and 48-hour weeks. The superiority of the 40- to the 48-hour week in this respect is probably due to the reduction of fatigue and boredom, greater concentration of effort, and improved morale of workers. It is not clearly known why the 36-hour week should yield less output per hour than the 40-hour week; the fact that the former involves a greater proportion of time spent in warming up and gives less opportunity to maintain skill through practice may account for its inferiority.

The total picture, therefore, seems to be that excessively long work-days throughout the week reduce total production because of increased accumulation of fatigue from day to day, liability to sickness and accidents, and unauthorized absenteeism; also, moderate reduction in the number of working hours to a point below that yielding maximum total production may result in a gain in hourly output. Thus, in deciding on the optimum number of hours of work, a choice must be made between obtaining maximum total production or maximum production per hour of work.

METHODS OF WORK

Changes in methods of work which may alleviate fatigue by reducing the effort expended in work include improvement in pace of work, technique of work, and equipment.

PACE OF WORK. The pace of work involves two considerations: (*a*) load, or the amount of work done per operation, and (*b*) rate, or the amount of work done per unit time. The significance of load is illustrated in a study of women doing continuous work (Cathcart et al., 1927). It was found in this instance that carrying loads exceeding 40 per cent of body-weight produced unduly high expenditure of energy. The influence of rate of work is shown in a study where a load of brick was transported by wheelbarrow at different speeds of walking (Crowden, 1928). The normal brisk walk involved a lower expenditure of energy than either slower or faster speeds. These and other findings indicate that for every job an optimum load and rate of work which make possible relatively high production at relatively low physiological cost to the worker. This optimum pace should be the one adopted when the object is to perform protracted work at maximum efficiency.

TECHNIQUE OF WORK. The way in which the individual does his work—deftly and with dispatch, or clumsily and laboriously—is an important determinant of fatigue. Efficient technique is largely a matter of economy of motion and ease of motion. In part, efficiency depends on the arrangement of materials and the work-pace. The most efficient arrangement will take into account comfortable posture, adequate support of the body and limbs, and freedom of motion. Materials and equipment should be organized so that standing, walking, bending, lifting, reaching, and other strains are reduced to a minimum.

Regarding economy and ease of motion the general objective is that the worker perform the minimum number of necessary movements that can be made with comfort. Myers (1925) has suggested six guiding principles that can be used in the development of greater efficiency in motion:

A. Successive movements should be arranged so that one passes easily into the next. The sequence should be so planned that few decisions or fluctuations of attention will be required.
B. Continuous curved movements are more likely to achieve the first principle than successive straight ones.
C. The number of movements should be reduced as far as is possible, to be compatible with comfort.
D. Whenever possible both hands should be used simultaneously.
E. No more effort should be exerted than is necessary to the satisfactory accomplishment of the work.
F. Materials and movements should be coordinated so that full advantage is made of momentum.

In addition, when the nature of the work permits, the performance should be made rhythmical, since this seems to reduce the necessity for hesitation and decision. Rhythm is thus beneficial in allaying fatigue and results in greater work output. Arrangements for rhythmical work should include adjustment of machines so that their rate and rhythm are coordinated to those of the worker. Auditory stimuli can be used with effect to guide work rhythms. Where music is used for this purpose, its tempo should, of course, be suited to the nature of the job.

EQUIPMENT. The use of tools and equipment greatly increases human productivity. However, working 8 hours with equipment is not necessarily less fatiguing than working 8 hours without it, unless the equipment is designed to effect real economy of effort. This means that equipment must be built to match the psychological characteristics of the operator.

One aspect of this problem concerns designing equipment with respect to human motion. Such factors as size, shape, location, distance of movement, and ease of movement of various units and controls must be taken into consideration. An Associated Press report (1943) indicates the benefits obtained from observing such factors in machine design. A typewriter with a new keyboard designed by Dvorak for use in the US Navy Department was found to improve output by about 35 per cent. An important feature of the new keyboard is that it distributes about 56 per cent of the labor to the right hand, which is dominant in most people, whereas the standard keyboard gives about 57 per cent of the work to the left hand.

A second group of factors with respect to which equipment should be designed concerns the perceptual characteristics of the worker. Illustrative of studies in this area are those which have been done on the problem of visibility and the recent wartime research on the optimum design of communications equipment relative to auditory requirements of the operators.

The third group of considerations in machine design is concerned with the various central or integrative characteristics of humans, including the attributes of such processes as attention, learning, transfer of training, and set. An example of this type of study is one intended to shed light on the habit interference that might occur when a pilot flies a plane that has different controls from those of the plane with which he is familiar. By using a mock control-panel under laboratory conditions, it was found that after the operator had learned one set of controls, a change in their positions on the panel had an adverse effect on his speed and accuracy of their operation (Weitz, 1946). Consistent identification of the controls by means of distinctive colors or shapes markedly reduced such effects, suggesting the advisability of coding airplane controls in a standard manner.

While gains in performance resulting from improvements in design of equipment are not to be attributed in large measure to fatigue reduction, such improvements are nevertheless consistent with the principle of minimizing undue effort and strain.

ENVIRONMENTAL FACTORS

Proper control of many corrigible factors in the working environment can contribute to the alleviation of fatigue. Environmental factors include noise, vibration, illumination, and atmospheric conditions.

NOISE. The results of scientific research generally agree with common experience in showing that the introduction of noise is initially detrimental to performance. However, the individual seems to be able to become accustomed to noise, so that after a time his performance regains its previous level (Morgan, 1916). This adaptation, however, is accomplished at the expense of additional strain and effort, including greater muscular tension (Morgan, 1916; Freeman, 1939) and increased metabolic rate (Laird, 1929; Harmon, 1933; Freeman, 1939). There is evidence that the excessive energy expenditure may disappear after habituation to the noise (Harmon, 1933; Freeman, 1939), but it is doubtful that this can occur when the noise is intermittent or very severe. In any event, exposure even to intense airplane noise for as long as seven hours did not seem to have any consistently deleterious effects on performance of a wide variety of tasks (Stevens, Egan, et al., 1946). Perhaps the main reason for reducing common noises in the working environment is their irritating and unpleasant effects on the workers (Bartlett, 1934; Stevens, Egan et al., 1946).

Noise can be reduced through the use of soundproofed rooms, sound-absorbing surfaces on walls and ceilings, mufflers, partitions separating different parts of the workroom, and other devices for reducing machine noises, by isolating noisy operations, and by supplying workers with ear defenders or plugs. Civic and factory educational campaigns against unnecessary noise-making may also be helpful.

VIBRATION. Vibration of the floor or chair frequently occurs when machines are running near-by. Burtt (1929) points out that this vibration may force one continually to make small, fatiguing muscular adjustments to counteract it. In one factory, where vibrations were transmitted to sewing-machine operators, each leg of the worker's chair was placed on a spring, and felt pads were used as footrests. In this way, the vibration was lessened greatly and fatigue was considerably reduced.

ILLUMINATION AND ATMOSPHERIC CONDITIONS. Inefficient illumination contributes to fatigue by increasing the strain and discomfort of work. The factors involved in proper illumination and vision include sufficient brightness, distribution of light which minimizes glare and contrast effects, elimination of flicker, and absence of visual defects in the worker.

Unsatisfactory atmospheric conditions augment the discomfort resulting from work, and in extreme cases may interfere with proper physiological functioning. Atmospheric factors that are significant in affecting performance include temperature, humidity, air circulation, and supply of oxygen.

PERSONAL FACTORS

A number of personal habits and characteristics of the individual influence his fatigability. Clothing which is comfortable and does not restrict freedom of movement should be worn. Its weight should be suited to the surrounding atmospheric conditions. The importance of proper habits of rest and sleep has already been made evident. Also, low resistance to fatigue may be associated with poor health. People who tire rapidly and are lethargic should undergo a medical examination to check on possible organic defects.

DIET AND DRUGS. Proper diet is important, not only because it is necessary to health, but also because it provides fuel for action. The more strenuous one's daily activity, the relatively more food he must consume if he is to maintain his health, weight, and maximum production (Kraut and Muller, 1946). Crowden (1932) estimates that had it been necessary during World War I to halve the bread ration of

British workers, industrial production would have suffered seriously.

Adjustments in the nature of the diet may aid in the alleviation of fatigue. Strenuous work in hot, dry climates frequently results in heat exhaustion because of the excessive loss of salt through perspiration. This loss can be counteracted by supplementing the diet with additional salt (NRC Committee on Work in Industry, 1941). There have been claims that the addition of certain foods to the diet is helpful in delaying the onset of fatigue. Included among the substances to which such action has been attributed are gelatin and various vitamins (*e.g.*, Kaczmarek, 1940; McCormick, 1940). However, other investigations (*e.g.*, King, McCaleb, et al., 1942; Keys and Henschel, 1942) have not corroborated such findings, so that their general pertinence remains in question. In any event, deficiency in vitamins, proteins, or other essential food elements must be avoided if vigor and resistance to fatigue are to be at maximum.

The distribution of food intake has an influence on working efficiency. It has been shown that industrial efficiency may be increased when the ingestion of the usual amount of food is divided between five meals a day rather than the usual three (Haggard and Greenberg, 1935). The improvement probably results in part from the value of the food as fuel. It is also perhaps a reflection of the fact that the metabolic cost of work tends to be less after a light meal than after a heavy one (Hovland, 1935).

Certain drugs have value in counteracting fatigue. The influence of caffeine, the principal drug in coffee and tea, in counteracting sleepiness and as a general stimulant has long been recognized. Benzedrine sulfate and ephedrine hydrochloride have been shown to relieve feelings of fatigue and boredom. Benzedrine has also been shown to be useful in offsetting fatigue decrements in performance. However, benzedrine doses exceeding 20 milligrams produce a temporary decline in psychomotor and thinking performances (Carl and Turner, 1939), pointing to the need for careful administration of the drug. Recently, still another drug, Δ-5 pregnenolone, has been found to reduce the output decrement in motor work (Pincus and Hoagland, 1944). Such drugs as benzedrine, ephedrine, and pregnenolone should not be used without expert pharmacological supervision. Nevertheless, when properly used, they do provide us with another possible tool for alleviating fatigue, at least in emergency situations.

Emotional Factors. Feelings of fatigue and deterioration of performance are often associated with a disturbed emotional state of the worker. A clinical study of industrial workers disclosed that fatigue and loss of sleep were prominently associated with the low emotional states during which workers frequently had accidents (Hersey, 1936). The interrelation of fatigue and emotional factors is further demonstrated in a war time study of fatigue among student pilots (Graybiel, Horwitz, and Gates, 1944). Overfatigue, i.e., fatigue not relieved by a good night's sleep, was found to be associated with such conditions as long hours of duty, lack of sleep, anxiety over dangerous work, and unsatisfactory personal relations with instructors.

The basis of the relationship between emotions and fatigue is not clearly understood. It is known, of course, that both are complex psychosomatic states. Emotions like fear and anger involve widespread organic changes which are conducive to efficiency of strenuous motor activity, and which bear some qualitative similarity to the physiological changes involved in work. Thus it has been suggested that prolonged emotion results in a condition characterized by physiological and conscious elements similar to those resulting from prolonged work (Sadler, 1929). It is likewise possible that prolonged work results in a condition sufficiently similar to emotion that emotional reactions will occur on slighter provocation than that required when the individual is not fatigued.

The evidence points to the fact that a program for the alleviation of fatigue should include steps for helping the worker to adjust emotionally to his job. Conversely, the apparent adverse influence of excessive work on the emotional condition of the worker adds another important reason for such programs.

READING REFERENCES

Bills, A. G., *The Psychology of Efficiency*. New York: Harper & Brothers, 1943, pp. 1–167; 199–232.

Huxtable, Z. L., M. H. White, and M. A. McCartor, "A Re-performance and Re-interpretation of the Arai Experiment in Mental Fatigue with Three Subjects." *Psychological Monographs*, 1946, **59**, No. 275.

Maier, N. R. F., *Psychology in Industry*. Boston: Houghton Mifflin Company, 1946, Chap. 14, "Fatigue"; Chap. 15, "Psychological Fatigue and Related Phenomena."

Poffenberger, A. T., *Principles of Applied Psychology*. New York: Appleton-Century-Croft, Inc., 1942, Chap.

6, "Work and Rest"; Chap. 7, "The Effects of Distraction upon Achievement: Noise"; Chap. 23, "Reducing the Cost of Work."

Robinson, E. S., "Work of the Integrated Organism," in C. Murchison (ed.), *A Handbook of General Experimental Psychology.* Worcester: Clark University Press, 1934, pp. 571–650.

Ryan, T. A., *Work and Effort; the Psychology of Production.* New York: Ronald Press Co., 1947, pp. 20–208.

Viteles, M. S., *Industrial Psychology.* New York: W. W. Norton & Company, Inc., 1932, Chap. 21, "Industrial Fatigue"; Chap. 22, "The Elimination of Unnecessary Fatigue."

13. THE PROCESS CONCEPTION OF PSYCHOLOGICAL ILLNESS

MATTHEW N. CHAPPELL, PH.D.

The Psychological Corporation. Formerly: Department of Psychology, Columbia University.

It is sometimes charged that, for the understanding of illnesses associated with emotion, American academic psychology is "sterile." The charge seems ill founded. Rather, it may be that the nature, origin, and control of those discomforts commonly called anxiety neuroses, neurasthenia, and, more generally, psychosomatic disorders may be understood more fully in terms of activities and processes known to occur in organisms—learning, forgetting, emotion, and the effect of summation of stimuli—than in more literary terms.

The basic psychological considerations required to understand the development and control of illness of psychological origin, and the orientation of these discomforts with regard to illness in general, will be sketched here within the framework of the physical view of nature and the basic organic processes.

COMMONNESS OF MANIFESTATIONS

Earlier accounts of discomforts of psychological origin stressed the individuality of each case. Medical and psychological examinations of large numbers of emotionally overactive men and women seen in clinical practice, as well as in connection with experimental and clinical studies of the relation of psychological factors to peptic ulcer, pseudo-peptic ulcer, arthritis, anxiety neurosis, constipation, and neurasthenic pains, indicate a marked commonness both in manifestations and in psychological behavior.

The following six conditions were found to be common:

1. Digestive disturbance, varying in degree from loss of appetite to spasm of the pyloris and severe pains in the epigastrium.
2. Fatigue, mild or severe.
3. Diminished sexual activity, which the subject may attribute to disinterest caused by fatigue or which may sometimes amount to impotence in males.
4. Proneness to fear, varying from specific situations (fear of certain foods) to an ever-present and most harrowing fear of being afraid.
5. Proneness to anger, manifested only in the relative security of the home or over a wide range of situations.
6. Excessive preoccupation with unpleasant thoughts which follow each other in an unbroken chain. The thought content concerns troubles, futility, failure, calamity, catastrophe, and disaster.

It can be recognized immediately that these conditions would be expected to arise automatically from prolonged and strong emotional activity.

The gastro-intestinal tract goes into hypermotility and spasm in emotion, the fundus becomes flaccid and the blood supply to the pyloris is diminished. Loss of appetite and tissue irritations become understandable. Energy is released to the blood stream rapidly in emotion and muscular tension mounts. Fatigue must result and be related directly to the intensity and duration of the emotion. The blood supply to the splanchnic area is restricted in unpleasant emotion and feelings of zest in the muscles are replaced by fatigue. Decreased sexual interest, or even impotence in some males, follows.

The fears and preoccupation with dire thoughts are readily understandable in terms of the laws of association and the volume of nervous impulses delivered from the viscera to the brain which controls associations. The commonness in the physiological changes which constitute fear and anger, and the ease with which humans shift from one to the other, would account for undue irascibility when marked fears are present.

RELATION OF PSYCHOLOGICAL ILLNESS TO ILLNESS IN GENERAL

For purposes of clarity consider briefly the specific case of a woman thirty-five years of age, who had strong pacifist leanings. After her husband joined the Army she suffered greatly from loneliness, was in conflict over her pacifism ideas and the needs of the war, and worried incessantly about herself and her husband. She developed a loss of appetite. She also had frequent panic reactions apparently without adequate cause, acquired many excessive fears including the belief that she was going insane, was impatient and irascible with her two-year-old son, and was convinced that her husband joined the Army to get away from her.

Similar cases are not uncommon. They are probably the most grossly mishandled of illnesses—by psychologists, psychiatrists, neurologists, and physicians alike.

In order to gain proper *perspective* on the problems of illness caused by psychological processes, it is well to consider Pike's formula (1935) for the maintenance of life:

$$L = f[(E)(PC)(G)(N)].$$

Life (L) is a *function* of the environment (E), the physicochemical basis of the organism (PC), the glandular system (G), and the nervous system (N). As stated, the four variables are dependent. A change in one of them necessitates a change within one or more of the others if Life is to remain relatively constant. Under this view, the older distinctions between organic and functional begin to weaken, for no change can occur in the physicochemical conditions without affecting one or all of the remaining conditions. Similarly, it appears that all disorders may be psychosomatic, and the word becomes no more meaningful than "illness." Further, it indicates that similar *manifestations of disturbance* might be expected to arise from *different primary causes*; that, for example, conditions diagnosed superficially as *"anxiety neurosis"* might be found on exhaustive examination to be physicochemical or glandular in origin; that the means to effective correction depends upon the variable concerned *primarily*.

Reconsider the case given above in the light of Pike's formula. All of the manifestations, including the psychological behavior, could arise from infection or presence of foreign proteins in the body, glandular disturbance, pathology, prolonged exertion, or from psychological processes. There is no way of knowing from the data given whether psychological factors are primary or merely secondary results of primary changes in physicochemical or glandular conditions. But determination of the *primary* cause is essential to rapid correction of the disorder.

The pious structure of literary psychology that might have been developed in the above case is obvious even from the brief description. The fact is, however, that this woman was completely symptom free three weeks after her infected tonsils were removed.

As appears from Pike's formula, no illness can be declared to be of psychological origin with certainty before exhaustive physical and physiological tests are made to determine that it is not physicochemical or glandular in origin. Only when such tests have been made is there justification for *assuming* that the condition *may be* primarily psychological.

PSYCHOLOGICAL ILLNESS IN TERMS OF LEARNING AND FORGETTING

Assuming that a discomfort is of psychological origin, how can we account for it in terms of the processes known to academic psychology?

The basic characteristic of these people is that they become emotional more easily and more frequently than is normal to them or to other people. They are able to produce an emotional reaction of major magnitude in response to a greatly reduced cue. From this point of view the excessive emotional activity appears as a skill, an unfortunate one, perhaps.

The acquisition of a skill requires practice. The more it is practiced the greater the degree of ease with which the reaction is evoked. All learned responses and skills are subject to forgetting. The only way to prevent loss through forgetting from manifesting itself is through added practice. Forgetting also sets the limit of perfection that can be achieved. A physiological limit is reached at some level. There the learning and forgetting processes balance each other no matter how much time is devoted to practice.

The acquisition of skill in becoming emotional seems understandable in terms of what the afflicted individual is doing. He is worrying long hours each day—dwelling on one thought after another, each of which evokes or practices the emotional response.

In this situation, we have an excellent example of a "feed-back" system. The more the emotional response is exercised, the stronger the response becomes and the greater the volume of impulses it sends back to the brain. These impulses determine, in part, what thoughts can next be evoked readily. By the time the emotional skill is well developed, it exerts such influence on associations that the worrier has difficulty in evoking any thoughts that are not unpleasant. At such a time, he may be practicing his emotional skill most of his waking hours. Forgetting gets no opportunity to rid him of the skill because he practices too incessantly.

Eventually he arrives at a level of perfection in which he can produce the "panic" reaction in the absence of any apparently adequate cause. In this reaction his heart races, his gastrointestinal tract goes into hypermotility and spasm, his throat constricts, his limbs feel powerless, and he expects to die on the spot. This is the skill at the peak of its perfection. It is the fear reaction at its maximum magnitude. When it occurs in the absence of any known organic, glandular, or physico-chemical disorder, it is commonly said to manifest an "anxiety neurosis." It may be more meaningful to call persons so afflicted merely expert worriers.

In addition to thinking incessantly about troubles, fears, etc., the person who has socially acceptable manifestations of psychological illness, for example, stomach trouble, gains considerable practice through embracing every possible opportunity to recite his woes. Recitation, as we know, is one of the more effective learning procedures.

CORRECTION OF ILLS OF PSYCHOLOGICAL ORIGIN

If the above approaches an accurate evaluation of the basic processes involved in illness of psychological origin, the needs for correction, in terms of academic psychology, are clear. Forgetting will obliterate any skill, and the more highly developed the skill is, the greater will be the absolute loss in unit time in the absence of practice. Further, insight into the nature of his difficulty will relieve the expert worrier of many causes for his worry.

In addition to furnishing insight, it would seem necessary only to arrange conditions so that all forms of the emotional practice are stopped. It is not difficult to stop the practice obtained by recitation, but stopping the silent review presents greater difficulty. The circle between the viscera and the thoughts is strong in these cases. Perhaps the easiest way to break through it is by planned and deliberate attention to a selected category of recollections, such as thinking about carefree childhood experiences each time one finds himself worrying.

No great skill in turning the attention to more pleasant thoughts would be expected in the early stages, but with constant practice, the worrier reaches a proficiency at which he can keep his attention off himself for minutes at a time. This interferes with and reduces his emotional practice and he begins to lose part of his skill. The "feed-back" mechanism now begins to work in reverse. As the emotional skill decreases, it is easier to break through the circle of unpleasant emotional thoughts. The interruptions are more frequent and longer, and the loss through forgetting becomes greater. Eventually, the emotional level would be expected to approach normal and all of the manifestations of high emotional activity disappear of necessity.

Evidence from experimental studies of peptic ulcer and pseudo-peptic ulcer, using control and experimental groups, clinical studies of osteo-arthritis and rheumatoid arthritis, anxiety neurosis, constipation, and neurasthenic pains, indicate that the application of principles of academic psychology may be made with results highly gratifying to those afflicted with illness in which psychological factors play an important part.

There is no available evidence, experimental or otherwise, to show that peptic ulcer or arthritis is *caused originally* by psychological conditions. The evidence shows only that after people have had these disorders for a long time, they recover more readily when the emotional level is reduced by psychological means.

READING REFERENCES

Chappell, M. N., *In the Name of Common Sense; Worry and its Control.* New York: The Macmillan Company, 1938.

———, and H. F. Kramer, "Constipation—A Training Problem." *Medical Times,* June, 1937.

———, and T. I. Stevenson, "Group Psychological Training in Some Organic Conditions." *Mental Hygiene,* 1936, **20,** 588–597.

Chappell, M. N.; J. J. Stefano; J. S. Rogerson; and F. H. Pike, "The Value of Group Psychological Procedures in the Treatment of Peptic Ulcer." *American Journal of Digestive Diseases and Nutrition,* 1937, **3,** 813–817.

Pike, F. H., "A Naturalistic Definition of Life." *Anatomical Record,* 1935, **64,** Supplement **1.**

14. DRUGS AND SMOKING

ST. CLAIR A. SWITZER, PH.D.

Professor of Psychology, Miami University; Registered Pharmacist (Michigan). Formerly: Assistant Psychologist, US Public Health Service.

Excessive ingestion of any chemical substance is harmful to the human organism, the extent of the harm varying with the physical constitution of the individual. Moderate use of certain so-called harmful drugs may be beneficial psychologically. The difficulty in generalizing concerning the physical effects of drugs is illustrated by the division of opinion among members of the medical profession regarding the effects of smoking. Certain physicians hold that many cardiac disorders can be traced directly to habitual use of tobacco, while others insist that there is little or no evidence to support such a view. But whether or not organic heart ailments can be attributed to smoking, many individuals believe their hearts have been affected in this way. Because the behavior patterns of such individuals are often modified by the belief just as completely as they would be by a demonstrable organic defect, it is important that we examine certain psychological factors that may be associated with smoking and the use of drugs.

TOBACCO

Hull (1924) showed that immediate mental and physiological influences of smoking tobacco were small. Efficiency in such mental activities as reading, mental addition, rote learning, and the like, was altered very little or not at all; but some disturbance of the more delicate muscular coordinations was demonstrated. Haggard and Greenberg (1934) found that smoking produces an increase in the discharge of adrenalin into the blood stream, releasing blood sugar from the liver and producing a feeling of temporary well-being in the smoker. This effect disappears within less than an hour after smoking.

Without more extensive data regarding such variables as general health, degree of tolerance, and physical habits, it would be foolhardy to attempt an evaluation of the physiological results of smoking in individual cases, but certain psychological assets and liabilities of the habit may be noted.

On the positive side tobacco-smoking furnishes visual, olfactory, and gustatory sensations which are pleasurable. Beneficial mental repose and release from tension during and after smoking are generally reported, but it is doubtful whether these values result from the physiological effects of the tobacco. The rhythmical processes of cigarette, pipe, and cigar smoking are conditioned elements in the relaxation patterns of millions of human beings and it is safe to venture that suggestion and habit are more important than the drug itself in the production of the principal pleasurable effects that accrue from smoking (although this is disputed by some physiologists, for example Haggard and Greenberg, 1934).

On the negative side smoking frequently becomes an unhandy and sometimes troublesome compulsion. It is during periods of emotional disturbance that smokers lean most heavily on "the weed." This holds true both for elation and despondency, but it is more pronounced during the latter state. Intervals of nervous excitement and nervous exhaustion call for some kind of escape and smoking appears to provide such an escape. In some instances individuals experience considerable mental unrest and worry because of the fear that they have or may become slaves to the tobacco habit. Any behavior pattern that results in such compulsions and conflicts is undesirable. The problem of balancing the psychological weights for and against tobacco-smoking belongs primarily in the field of personality adjustment.

CAFFEINE

Coffee, the "cola" soft drinks containing caffeine, and tea which contains a similar drug, are the beverages most universally used. Physiologically, caffeine is an interesting drug because of its apparent nonconformity with the tolerance principle. Studies show that small and medium-sized doses are stimulating, but large doses may be somewhat depressing. Speed of performance improves on most psychomotor tasks but hand steadiness decreases under caffeine influence. Individual reactions to the drug appear to change very little with use. Psychologically, the drinking of caffeine beverages has much in common with tobacco-smoking. Habit plays an important role in the effects produced. Many people think they must have their caffeine drink at certain regular intervals and to be deprived of it produces psychological unrest. The social factor no doubt contributes to the stimulating effect which habitual users value so highly.

ALCOHOL

Alcohol has a depressant effect upon all psychological functions and is a general efficiency reducer. Most people think of alcohol as a stimulant but all experimental evidence is in agreement regarding its depressant action on the functions of the central nervous system (Jellinek, et al, 1945). The degree of susceptibility to its influence and the speed of its reaction are related to numerous factors, such as body weight, amount of tolerance that has been established, and the dietary habits of the user. Recent research indicates that certain elements in the vitamin B_1 complex serve in neutralizing or minimizing certain physiological effects of alcohol. It is suggested that this dietary factor may account in part for periodic fluctuations in the alcoholic tolerance of a given individual.

Alcohol frequently affects human behavior patterns in the reverse order of their acquisition. In the normal course of development we learn to walk, then to talk, then to reason, then to use judgment, and finally to accept responsibility. Responsibility from fears and worries may be lost first by the person under the influence of alcoholic intoxication, then judgment may be affected, then reasoning become confused, then talking blurred, and finally the power of locomotion may be disturbed and lost. But this pattern is not a consistent one among all drinkers, or even for the same drinker on all occasions.

Excessive use of alcohol was once accepted as the cause of many marital maladjustments and unhappy homes, but clinical evidence indicates the reverse, that marital maladjustments and unhappy homelife are often the cause of excessive drinking. As the psychoanalysts point out, alcohol furnishes one of the easiest psychological escape devices available to man. Used in moderate quantities by reasonably well adjusted persons, alcoholic beverages possess relaxing and socializing properties which may very well overbalance any deleterious physiological effects. But the abnormal or unstable personality is likely to find in alcoholic excess the solution of all problems. Therefore, he faces the danger of dipsomania, a condition which is usually far worse than the psychological disturbance that led to it. The psychology and treatment of alcoholism have been thoroughly examined by Strecker and Chambers (1938), and a critical review of the literature on alcohol has been prepared by Marshall (1941)

OPIUM, MARIHUANA, AND COCAINE

THE OPIUM FAMILY

Opium, morphine, and heroin are depressant, pain-relieving, habit-forming drugs, of which heroin is the most treacherous in its action. Addiction to an opiate may result from a variety of causes. The habit may be formed during a period of severe physical illness when the drug is administered on medical prescription, but such cases are rare. More often the habit is acquired more or less deliberately when the drug is secured illegally and is used to escape from reality, and as a relief from mental conflict. Among the well-to-do, addiction has not infrequently resulted from the desire to experience a new thrill. Once the habit is acquired, the victim becomes a potential menace to society because he will go to unusual lengths to satisfy his uncontrollable craving for the drug.

In the realm of opiate or narcotic usage, two behavior principles have received wide acceptance. While there are exceptions which render these rules invalid in individual cases, they are well supported by clinical evidence: (1) Prevention of addiction through moderate usage is impracticable. (2) "Once an addict, always an addict." In nonaddicts the depressant action of the opiates produces a lowering of efficiency of mental and motor functions. Administration of opiates to addicts, however, generally results in an increase in efficiency of these functions.

MARIHUANA

The smoking of marihuana has become widespread in the United States, and extreme police vigilance in locating and destroying the plant is exercised in certain areas. Cannabine, the active principle of marihuana, affects the higher nerve centers in several ways. The first effect is a feeling of great exaltation accompanied by an illusory enhancement of physical and mental powers. Everything seems possible. Space and time orientations are distorted. All sounds appear to be magnified. This first stage is followed by sensations of muscular weakness which usually result in sleep. Continued use of marihuana may lead to addiction and permanent personality changes. The addict experiences an increase in sexual excitement together with paranoid delusions and fits of delirious and irresponsible rage. This stage marks the onset of complete mental deterioration.

COCAINE

Cocaine is a stimulant for which the human body seems never to develop a tolerance. It releases the speech mechanisms and sharpens the mental processes temporarily. The user experiences an increase in energy and a feeling of expansive well-being while under its influence. This stimulation gives way later to a profound depression. The advanced stages of cocaine addiction resemble some forms of alcoholic disturbance in that cutaneous hallucinations are commonly experienced. General deterioration of the mental processes and demoralization result from continued use.

STRYCHNIN AND BENZEDRINE

These drugs and their salts are stimulants, and they provided increased mental and physical energy for a limited period. Strychnin and its salts are known to be dangerous even in small doses because of a powerful stimulating action on the central nervous system. Not so much is known about the general effects of benzedrine and its compounds but experimental evidence indicates that they improve performance on tests of steadiness, reaction time, tapping, handgrip, and maintained handgrip.

AUTOCOID SUBSTANCES

An often overlooked source of drugs is the human body itself where the endocrine glands produce a large number of chemical substances that have a profound effect upon human behavior. Some of these hormones or autocoid substances, such as adrenalin, thyroxin, and testosterone, have been produced in the laboratory. Adrenalin, which normally enters the blood stream from the adrenal glands during emergency situations of anger, fear, or the like, is the great energy-producing hormone. It is also administered therapeutically as a heart stimulant. It affects different personalities in markedly different ways, making some tearful, some joyful, and some fretful. Without it, however, the human organism would be a relatively uninspired and inefficient mechanism. Thyroxin, produced by the thyroid gland in the neck, appears to be an important catalyst, assisting in the process of elimination of body waste and thereby increasing physical and mental efficiency. A deficiency of thyroxin in the body produces a condition of mental and physical sluggishness, an excess causes nervous irritability, restlessness, and emotional instability. Deficiencies are overcome by thyroid therapy. Testosterone, a synthetic product, may enhance the sexual vigor of the male, and also effect some favorable personality changes. Research with human subjects has not been extensive and it is therefore too early to advance definite conclusions about its psychological properties.

Other important hormones are secreted by the ductless gland system. Their characteristics and functions are in doubt. They may be the cause of some very puzzling personality differences in efficiency and happy adjustment. Exact knowledge of their chemical structure and function may make possible the prevention and control of many forms of mental and physical maladjustment.

SUMMARY

This chapter presents a survey of the psychological effects of certain drugs which man ingests voluntarily or manufactures within his own body. While some of the drugs described may not be actually harmful, many of them are, and the harmless ones may produce no beneficial psychological results. All depressant drugs and the stronger stimulants are potentially dangerous from both a physiological and psychological standpoint. Some drugs have social or psychological values but drugs are most frequently taken voluntarily as a means of escape from conflict situations. Only among the milder stimulants are there a few drugs that the average human being can use with impunity over a prolonged period of time.

READING REFERENCES

Haggard, H. W., and L. A. Greenberg, "The Effects of Cigarette Smoking upon the Blood Sugar." *Science,* 1934, **79,** 165–166.

Hollingworth, H. L., "The Influence of Caffeine on Mental and Motor Efficiency." *Psychological Monographs,* 1912, No. 22.

Hull, C. L., "The Influence of Tobacco Smoking on Mental and Motor Efficiency." *Psychological Monographs,* 1924, **33,** No. 3.

Jellinek, E. M., et al., "Alcohol, Science and Society." *Quarterly Journal of Studies on Alcohol,* 1945.

Marshall, H., "Alcohol: A Critical Review of the Literature, 1929–1940." *Psychological Bulletin,* 1941, **38,** 193–217.

Strecker, E. A., and F. T. Chambers, *Alcohol: One Man's Meat—.* New York: The Macmillan Company, 1938.

Walton, R. P., *Marihuana.* Philadelphia: J. B. Lippincott Company, 1938.

CHAPTER III

THE INDIVIDUAL ADJUSTING TO HIS ENVIRONMENT

15. Motivation and Self-Direction	Douglas H. Fryer
16. Interest in Work and Play	Donald E. Super and Jack W. Dunlap
17. Aptitude and Intelligence	George K. Bennett
18. Personality and Physique	Otto Klineberg
19. Guided Learning	T. R. McConnell and G. Lester Anderson
20. Vocational and Educational Satisfaction	Robert Hoppock and Donald E. Super

15. MOTIVATION AND SELF-DIRECTION

DOUGLAS H. FRYER, PH.D.

Associate Professor of Psychology, New York University, and Secretary, Richardson, Bellows, Henry and Company, Inc.; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology. Formerly: Chairman, Department of Psychology, University College, New York University; Expert Consultant to the Secretary of War; Collaborator, US Forest Service; Chief, Recruiting, Selection and Classification Research Unit, Personnel Research Section, AGO, Department of the Army; Chief, Personnel Research Branch, Civilian Personnel Division, Office of the Secretary of War; Adjutant, Morale Branch, General Staff, U. S. Army; Director, Vocational Department, Central Y.M.C.A., Brooklyn, New York; President, American Association for Applied Psychology and Association of Consulting Psychology.

The psychological past of everyone lives on in him throughout his life. We forget this truism in dealing with other people. We know it from experience, but we forget it in trying to understand our own feelings. We think and behave in a manner which is predetermined by our past or by the psychological being we have become.

The influences of heredity and environment are interwoven in us as the warp and woof of our existence. What heredity contributes is vague but important. It lays its hand upon our energies and abilities; it answers the question finally of how far we shall go.

Our environment, on the other hand, answers the questions: what direction shall we take, shall we be friendly or unfriendly, for or against, cooperative or competitive, generous or stingy, approving or resentful, silent or talkative, and so on, in every situation in life. It does this by the manner in which conditioning and learning goes on in us as we live our lives in an environment formed of natural resources, economic conditions and social contacts.

MENTAL SET

Mental set is the name used to describe the psychological past affecting particular response situations in the individual. It is inclusive in meaning of such terms as attitudes, hungers, muscular tensions, organic states, emotional excitements, and everything that motivates us.

Every living organism is driven to activity by internal energy. The energy conditions, which are internal stimuli upon interoceptors and proprioceptors, cause man, or any other organism, to act upon its environment. Thus the organism comes in contact with environmental stimuli. And as it acts upon the environment, the environment acts upon it. The environmenal stimulus, affecting exteroceptors, blocks or removes, molds or represses, in a word, directs the expression of bodily energy upon the environment.

This is the process of conditioning and learning, which goes on through life. It is one of seeking and avoiding environmental stimuli, or of establishing problem-solving stimuli to direct the forces of internal energy. Incentives is a common term for these problem-solving stimuli. Mental sets are established to react only to certain incentives. Also, certain incentives may initiate a mental set for reaction or thinking, replacing in large measure the original causation of behavior by internal energy.

By way of illustration, the tissue needs in the young child cause contractions in the stomach, what is called hunger pangs. The child expresses internal energy upon his environment by kicking and crying. An attendant stimulates him with toys, rocking, talking, and so on. But none stop the expression of internal energy until a bottle of milk comes into view. At first the child did not stop crying until the milk flowed into his stomach. Now he smiles and reaches. Already an incentive is determining the mental set where only tissue needs caused activity in the child before. Let us take a long jump in the child's development. The child is now a worried business man. He arrives home dead-tired and says he is not hungry and he is going to bed without his dinner. A dinner of steaming roast beef is on the table. The sight of it causes hunger pangs and the man eats a big meal. There was a tissue need for food, but the direction it gave to behavior was obscured by other energy conditions of the body, muscular tensions or fatigue. An incentive took over the direction of motivation. The sight of the roast beef and all of the fixings established the mental set to eat.

What we do automatically, what we do without conflict and painful decision, our rapid speech and reading, that which comes as a flash in thinking, the fine techniques for which we are envied and the inhibitions that make us social beings, all of these are performed through mental sets. When we don't know something, have no opinion to express, do not know what to do, we are without mental sets. Our attitudes, theories, and beliefs are expressed because of mental sets. We are Americans, communists, democrats, gypsies, patriotic, aggressive, unselfish, immoral, bankers, politicians, and so on, because of mental sets. Life depends upon the drive of internal energy. But an efficient and satisfying life depends upon the direction given this energy in mental sets.

THE DRIVE OF INTERNAL ENERGY

The drive of internal energy has been investigated in rats and other animals (Tolman, 1932; Warden, 1931). One method is to measure the amount of electric current absorbed in the effort to secure an incentive. Female rats will stand the most punishment of this kind to be with their young. The relative order of motivation in rats and other animals is, first, the drive in females to be with their offspring and then, in order of strength, thirst, hunger, sex, and, finally, and much less positive, the general drive to activity. No similar measurement has been made with human beings. But observational studies of man's behavior under various conditions indicates that he would respond similarly in drive of internal energy.

PREPARED UNITS OF RESPONSE

Mental sets are organized in relation to each and every object and activity surrounding us. They are prepared units of response. The condition of readiness that enables us to brake a car down quickly to a slower speed as various objects cross our line of vision is an ever-changing set in relation to the specific incentives of our driving. So with the holding of a spoon by a child and the inhibition of excretion until he is on the "potty." Stamping a letter, writing a figure, doing a sum; walking or jumping, toeing in or out, speaking quickly or slowly—all the differences so evident between people—are specific sets of motor and verbal coordinations.

If a person is asked to perform a series of similar mathematical problems, each problem in the series is likely to be done according to the manner established for the performance of the first problem. This will be true even if we introduce a problem somewhere in the middle of the series that can be performed in the same manner but can be done far easier in another way. The same mental set continues in operation as long as it works, even though it is not the best one. We are psychologically blind to better methods for the solution of problems, unless incentives are so arranged that existing sets are rendered inadequate.

We sense what we have been trained to sense; we behave as we have been trained to behave. We do not know what we would sense if we were free agents registering all that might come to our senses. We do not know how we would act if the incentives of our life were different. We fight for our ideals, whether wrong or right, because of our sets. We are murderers, hobos, Babbitts, martyrs, etc., because of them. That we are this way accounts for our differences and similarities. Psychologically, life for the individual at any one moment is the sum total of mental sets directing response.

MOTOR SET

Set is often described in terms of muscular coordinations, as when a runner is on the mark waiting for the gun. Back in 1889 when psychology was mostly theory, Muller and Schumann (1889) compared the lifting of weights with a normal motor set and with one artificially established by lifting a heavier weight. The weight-lifter stood blindfolded in front of a table and alternately lifted, with opposite hands, a standard weight of 627 grams and comparison weights of 826, 876, and 926 grams. All weights were of the same size. The three comparison weights were judged to be heavier in almost all instances. Then a much heavier weight, 2,476 grams, was substituted for the comparison weights, which always was judged heavier, as would be expected, than the standard weight of 627 grams. After this, the three comparison weights were substituted for the 2,476-gram weight. They were now judged lighter or equal to the standard, but never heavier as they had been judged before. A new motor set had been established and the three comparison weights seemed lighter than before in comparison with the standard weight. Muller and Schumann found that the new motor set was strongest immediately following the practice with the heavier weight and rapidly died out, but that with practice on the following day, it was again effective. An individual's normal motor set for weight-lifting, of course, is the product of continuous practice with various weights.

PREPARATORY SET

When a worker is getting set to perform a task his muscles contract or expand in preparation for the work although there is no overt expression in effort or performance (Katzell, 1948). The curves of muscle activity during this period of preparation or of getting set rise gradually until the time comes for the expression of motor activity. Always they represent a condition of partial (in relation to the set during the work) and increasing muscle activity from the time intention has formed to do something until the work is performed. This muscle activity may subside, of course, if intention changes. Also, it is probable that the involved muscle activity will subside gradually upon completion or cessation of work. A preparatory setting of the muscles is necessary for performance before beginning any task.

PERSISTENCE OF SET

Persistence of set has been measured by a comparison of the recall of tasks performed to completion and tasks left uncompleted. Zeigarnik (1927) examined a large number of activities for the persistence of set, including tasks involving both language and manual work. The workers in the experiment were allowed to complete some of the tasks but they were interrupted before completing others. Later they were asked to recall the various tasks which they had performed. They recalled more tasks that they had been unable to finish—in fact, nearly twice as many. This was true with both adults and children.

Various experiments have been performed upon the persistence of set in a similar manner to that of Zeigarnik (Lewin, 1935; Pachauri, 1935). Tasks interrupted close to their completion are recalled better than those interrupted earlier, illustrating increased tension as a task is brought toward completion. When recall was delayed twenty-four hours, both the finished and unfinished tasks were remembered about equally well. When tasks were interrupted by asking a worker to help someone in an adjacent room, spontaneous resumption after returning took place in all cases if the period of the interruption was only a few minutes, but if it was longer, resumptions were less likely.

This phenomenon of the persistence of set is familiar to everyone, as is the manner in which one is able to forget and ignore work planned so insistently earlier. When various tasks were performed by university students after attending lectures for from 6 to 7 hours and by office clerks after a day's work, recall was definitely in favor of the completed tasks. Fatigue seemed to have resolved the sets which usually would have persisted. This is a familiar experience too.

GOALS OF SET

Knowledge by a worker that he is set to do a task is necessary to the existence of the set and to the continuation of the task for which he is set to perform. The set exists only so long as the worker understands what he is doing or where he is going. A shift in goal changes the set. Let attention be diverted from the goal and the performance lapses into inertia. Let a worker believe that the completion of a task is not wanted, and he will not resume work. Add to the importance of the goal, and the persistence of the set will be greater. Destroy the goal, and work ceases.

The dependence of set upon goal was clearly indicated in an experiment by Marrow (1938). First, Marrow established that the recall of unfinished tasks was 57 per cent better than for the finished tasks—where the goal was to complete them. Then he reversed the goal of completion by instructing the workers that they would be allowed to work at a task only until sufficient mastery had been attained. Thus objective completion of a task would not resolve the set, only interruption would do it. Tasks actually completed were recalled 34 per cent better than those stopped before completion. Also, Marrow increased the importance of the goal to complete the tasks as instructed by giving competitive information. When this information was encouraging the unfinished tasks were recalled 93 per cent better than those finished; and when it was discouraging, 71 per cent better. Increasing the importance of the goal increased the persistence of set.

INCENTIVES

Incentives have been referred to earlier as the problem-solving stimuli of the environment. They are the stimuli which give direction to motivation or the internal energy of the body as it is expressed upon the environment. A survey will follow, of the experimentally established facts concerning the action of incentives.

NOISE

Noise is called a distractor and it is popularly assumed that distractors have a detrimental effect on efficiency. But distractors that have been investigated, such as gongs and buzzers, affect efficiency very little (Bartlett, 1934; Morgan, 1916; Pollock and Bartlett, 1932). Generally speaking, there is a slight slowing of the rate of work when the distractor is introduced, followed by a definite increase which is maintained consistently during its presence and by a slowing to the normal rate of work upon its removal. Distractors may have different effects on individuals. For example, music over the radio may facilitate reading or work by some and not by others, and a speech on the radio may inhibit it.

The actual effect of noise on efficiency is determined by the attitude of the worker. Baker (1937) has demonstrated this. He instructed three separate groups of workers differently, in advance of doing the same work, that distractors had (*a*) a detrimental effect, (*b*) a facilitating effect, and (*c*) a detrimental followed by a facilitating effect. He found that measured efficiency during noise corresponded with the effect each group had been instructed to expect.

The worker's attitude can make any ordinary distractor seem far off and unrelated to his work. It will determine whether a noise will be an incentive to work harder, or whether the noise is an excuse for more leisurely work. The goal of the worker is responsible for any effect of noise on performance with the usual noises of our environment.

RHYTHM

Regularly timed distractors or rhythms may be more persistent stimuli than irregular noises. But they are incentives only when the worker makes a goal of them. People do not "fall into" rhythms unknown to themselves and workers cannot be inveigled into working faster by hearing the rhythm. A worker must adopt the timing of the rhythm as a goal if his working rate is to be directed by the rhythm (Fryer, 1934). With such a goal, workers can perform routine tasks in time with auditory rhythms at a very high rate for them; and they say it is easier and pleasanter to work at high speeds in this manner when objective timing takes over the responsibility for keeping up the pace (Fryer, 1937).

KNOWLEDGE OF RESULTS

Knowledge of results is frequently an incentive affecting the goal of performance. Goals of achievement, or levels of aspiration as they are occasionally called, are based quite largely upon past performance. Workers must be informed of the degrees of efficiency of past performance to establish realistic goals for future performance. Workers are highly consistent in establishing their goals at different times (Frank, 1935). They are fairly consistent in establishing their goals in different tasks (Gardner, 1940; Gould, 1938). A worker sets a goal in any new task according to past performance in similar tasks. He is likely to aspire high or low in various tasks according to his usual manner of setting goals (Frank, 1935).

The effect of knowledge of results on efficiency when comparing conditions where it is given and withheld, is usually considerable (Book and Nowell, 1922). All activities are affected by it. In a most fundamental form of behavior, reaction time, Johanson (1922) showed an increase in efficiency of 6 per cent when knowledge of results was given. The effect of this incentive depends upon the form, the clearness of detail, and the part of the performance in which the knowledge is gained. Knowledge of past performance influences the efficiency of both learning and routine work. A learner or a worker can investigate these things for himself. Goals which are based on this information have a progressive effect upon his efficiency.

SOCIAL INCENTIVES

The social situation in which one works is an influence upon the goals of one's achievement, as probably every student has observed when he studies in a study hall, and every worker has observed when the foreman is looking on, and every speaker, when he practices his speech before a mirror. Allport (1920) found that free association was faster among adults working in groups than when working alone. But arguments were slower. The more intellectual the task, the greater the social inhibition upon it. Travis (1928–29) found that stutterers were faster in free association when working alone, which shows that a social inhibition exists among people with this speech handicap. The social situation influences the goal of work and efficiency may be inhibited or facilitated by it.

COMPETITION

When knowledge of results acts as an incentive it creates a goal for self-competition, although there may be many differences in the definition of the competitive goal. When individuals compete against other individuals there are additional differences in competitive goals. When individuals cooperate together as a team or a group, to compete with other teams or groups, there are further differences in the definition of goal. Seldom do two individuals ever have the same competitive goal even though the competitive incentives are the same (May and Doob, 1937).

Maller (1929) compared individual competition and group competition (individual cooperation) among school children of all ages and found that individual competition definitely was the strongest incentive. But group competition also was a strong incentive for accomplishment. Others have found similar results with various ages among both educational and industrial workers.

A comparison of individual and group competition shows a difference in the persistence of the competitive set. In Maller's experiment, work for self was consistently high, and rose slightly over a considerable period, whereas work for the group was twice as variable and declined in amount with time. Evidently the goal of individual competition is more stable than that of cooperation, as well as of greater initial strength.

Wide individual differences in performance exist in competition or cooperation. Such differences are influenced by ability, but the goal also determines these differences. Hurlock (1928) found that older children in the same group with younger children did not respond to the incentive of competition, nor did superior children with inferior children, nor did boys in a group with girls. Leuba (1930–31) found that slower multipliers improved vastly more than faster multipliers in the same group. The first quarter of the class increased only 34 per cent in the competition, and the last quarter of the class increased 71 per cent.

The background sets of the individuals in the group are responsible for the goals of present sets. What one aspires to now and accomplishes is the result of what one has aspired to and accomplished in the past, and what one aspires to in the future is the result of what one aspires to and accomplishes now. Seeking competitive incentives provides goals of increased accomplishment in the development of individual efficiency.

REWARD AND PUNISHMENT; PRAISE AND BLAME

The incentives called reward and punishment, praise and blame, are effective according to their social implications. Any specific punishment will not have the same effect upon a girl as upon a boy, or upon an employee as upon a prisoner. Reward is said to be a superior incentive to punishment in guiding youth; and praise in comparison with reproof. Yet there are many human situations in which this is not true, and in experiments among animals the greatest reward will not produce as great activity as can be secured through punishment.

Punishment increases the effectiveness of behavior either through the information given by it or as an annoyance to be avoided. Gilbert (1936) showed that learners learned better with electric shock than without it, and Johanson (1922) showed that punishment with electric shock lowered reaction time 15 per cent, which is more than twice as much as was caused by knowledge of results. Slight electric shock was found by Vaughn and Disern (1930) to be more effective in learning a maze pattern where none, slight, moderate, and intense electric shocks were used. Slight punishment calls emphatic attention to the necessity of benefiting from past errors. Stronger punishment affects individuals differently because of the varying implications of the punishment to the worker and it may or may not be an incentive to increased efficiency.

Rewards have been studied for their effect in education (Thorndike, 1933) and in animal psychology (Tolman, 1930) for their implications in understanding motivation. Bonus plans are compared in industry to see what rewards are more acceptable to the worker (Lytle, 1929). Extensive research effort has been expended in study of rewards and their influence upon performance in all branches of psychology. The reward offered, whether it be a bonus for the industrial worker, sunflower seed for the rat, a gold star for the school child, a college degree for the student, is an incentive only as it develops in the individual a goal of accomplishment.

Individuals choose their incentives according to mental sets of the past. They can establish rewards for themselves for increased accomplishment along with punishments to be avoided by this accomplishment. A college student can give himself a day off for certain attainment, thus making this performance his goal. A businessman can go home for the weekend without his brief case as a reward for intensive effort throughout the week or he can take it home as punishment for things left undone. Rewards and punishments are incentives easily manipulated by the individual.

Various investigators have shown that both praise and reproof have a significant effect upon performance. Immediate performance does not improve more from praise than from reproof. But the long-period effect of praise is far greater (Hurlock, 1924). Eaton (1933) has reported an interesting series of observations in which praise decreased steadiness in men and increased it in women. It was decreased in men in an equal amount to that caused by a short period of exercise in basketball. Reproof, on the other hand, increased steadiness in men and decreased it in women. This sex difference probably arises from differences in social training. Praise and reproof are subject to many shades in social definition and they affect people differently, even more than do rewards and punishments.

SUMMARY

The same incentives do not have the same effect on all individuals. Previous sets determine the influence of any new incentives. But group averages in experiments for various situations in which reward, punishment, distraction, competition, cooperation, and so on, are present imply varying degrees of universality in motivation.

If this is understood, individuals trying to help themselves to do better work, or teachers trying to encourage their pupils, or fathers trying to direct their sons, or foremen trying to stimulate workers, will seek first an acceptance of an incentive as a goal, before expecting it to effect performance. Any stimulus, to become an incentive, must fit in with the objective of a worker. While there are many incentives of fairly general acceptance by people, the individual chooses those that help him according to his psychological past; as they facilitate achievement of goals already established in his life.

INTENTION

The formation or definition of a goal by a worker has been studied as intention. The degree to which one's intention determines what one will observe or ignore, or what one will remember or forget has been indicated in various experiments. Kulpe (1904) investigated the superiority of a specific goal over a general goal in the learning of groups of nonsense syllables. Different specific goals, such as observing the number of the letters and observing the colors and positions of the letters, were used in the learning. Kulpe found when learning was according to these conditions that considerably more was learned of the things the learner wanted to learn and just as much of everything else. Moore (1936) has illustrated the importance of clarity of intention. The material to be learned in one experiment was a prose selection which consisted of one hundred and sixty-six words and sixty-seven ideas. One group took the ideas as its goal for the learning, a second counted the words, and a third, the control group, read the passage as an assignment in English. All were tested for recall of ideas. The first group scored 13 per cent higher than the control group and 630 per cent higher than the second group, thus illustrating the importance of intention in study as well as the superiority of a correctly defined goal for learning details.

To study without knowing for what he is studying; to read a book without knowing for what the information he reads will be used, usually gives the student or worker some valuable information. It is better than looking for the wrong things when reading or studying. If one does not know for just what he is preparing, this is perhaps the best way to study or work, that is, to study with a general goal to learn all he can. But to read or study in search of information that will be useful, that will result in a clear understanding of the goal or thing he wants to accomplish, there lies the greatest efficiency for the student.

Efficiency in manual performance is related to intention in similar manner. Let us set up an experiment (Fryer, 1937) in which we disturb the worker's intention to perform at an unchanging pace. He can perform, if undisturbed, at almost exactly the same rate minute following minute. But as the distracting stimuli are introduced, the worker adopts goals of performance to counteract their expected influence. Performance changes exactly in relation to the specific goals. Such work is guided pretty accurately by intention. But the worker has lost touch with the original rate, and performance is increasingly variable as long as the disturbing stimuli are present.

Almost anyone can raise his performance rate. There is no limit, it would seem, to what the individual can do provided the goal can be made real in his intention. The continuous breaking of athletic records illustrates this. Studies of motivation support this conclusion that ability can have unlimited expression if realistic goals are provided.

FRUSTRATION

The experiment just described illustrates a certain amount of frustration of intention but not more than is usual in many occupations. But when goals are adopted which are impossible or unreal, frustration is likely, resulting in uncontrolled anger, psychomotor tension, depression, remorse, embarrassment, and sometimes withdrawal from reality and a complete change of personality.

Most of us do not take flight from an oppressing environment, but we daydream of having the last withering word in a conflict of ideas, or we dream of childhood scenes when life was easy, or of realization through our children of what we cannot attain for ourselves. We sing songs that "the best things in life are free" and "I want what I want when I want it," but we recognize limitations on our self-indulgence. We are forced to direct most of our thinking and action in relation to our surroundings. The more adjustive our mental sets and their goals, the better we are able to face a world that is in opposition to too complete mental independence. For many of us, the problem of self-direction is how to avoid the dilemma of frustration and its resulting maladjustment. To do this, mental sets must be organized with realizable goals which achieve a satisfactory release from the driving energy within us and which gain for us a physical and mental adjustment.

SELF-DIRECTION IN SOCIAL CONTACT

Self-direction comes in the process of choice of incentives. Each environmental stimulus is tested and selected to become or not to become an incentive for future action. It is investigated by the individual to see how it satisfies his need. It is established as something in the environment to be sought or avoided. This applies to all material things and to all social contacts.

Incentives are effective generally with people of the same likemindedness, that is, when they have the same educational, occupational, and social background. Social groups are formed of people who are affected in like manner by incentives. Individual members of social groups change and develop as they are influenced by these incentives. They become more social or more antisocial as the incentives which make for social groupings are acceptable or unacceptable to them.

Dealing with social stimuli has become more important to the individual as society has become more complex. To respond correctly in business, political, or cultural groups is one of the fundamental goals of the individual today. Considering only the individual and not the organization providing social training, there is continually necessary a choice of social stimuli: of schools, friends, business associates, and so on. In this choice the psychological past of the individual exerts a strong influence. But in this choice changes are made in mental set to improve adjustment, avoid frustration, and realize social satisfaction.

READING REFERENCES

Bartlett, F. C., *The Problem of Noise*. New York: The Macmillan Company, 1934.

Diserens, C. M., and J. Vaughn, "The Experimental Psychology of Motivation." *Psychological Bulletin,* 1931, **28**, 15–65.

Lewin, Kurt, *A Dynamic Theory of Personality*. New York: McGraw-Hill Book Company, Inc., 1935.

May, M. A., and L. W. Doob, *Competition and Co-operation*. New York: Social Science Research Council, Bulletin No. 25, 1937.

Moore, J. H., "The Role of the Determining Tendencies in Learning." *American Journal of Psychology,* 1936, **48**, 559–571.

Young, P. T., *Motivation of Behavior*. New York: John Wiley & Sons, Inc., 1936.

16. INTEREST IN WORK AND PLAY

DONALD E. SUPER, M.A. (OXON.), PH.D.

Associate Professor of Education, Teachers College, Columbia University; Consulting Psychologist, American Institute for Research; Diplomate in Counseling and Guidance, American Board of Examiners in Professional Psychology; Certified Clinical Psychologist (New York). Formerly: President, Division of Consulting Psychology, American Psychological Association; Major, Air Corps, Chief, Psychological Branch, AAF Regional and Convalescent Hospital, Miami, Florida, and Research Officer, Psychological Research Unit No. 1, AAF Training Command; Associate Professor of Educational Psychology and Director of the Personnel Bureau, Clark University; and Director, Cleveland Guidance Service, Cleveland, Ohio.

JACK W. DUNLAP, PH.D.

President, Dunlap, and Associates, Inc., New York; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology. Formerly: Commander, USNR and Officer in Charge Aviation Free Gunnery Training and Research Unit; Director of Research, Committee on Selection and Training of Aircraft Pilots, National Research Council; Associate Professor of Educational Psychology, University of Rochester and Fordham University; President, Psychometric Society; New York State Association of Psychology; Division of Consulting Psychology, American Psychological Association; and Division of Psychology, New York Academy of Science; Editor, *Journal of Educational Psychology* and *Personnel Journal.*

A person's interests are the things to which he responds with a feeling of pleasure, the activities in which he engages or to which he reacts with satisfaction when observing or thinking of them. An interest may be specific, limited in its scope and perhaps in its duration; or it may be broad, basic, and enduring. Viewed operationally, interest is a response of liking. Considerable confusion has arisen in the past because of failure to specify the sense in which the term interest was being used, but despite this there has long been considerable emphasis on the importance of interest in education, work, and play.

As early as the seventeenth century the educator Comenius pointed out the necessity of making subject matter interesting. Later Rousseau, and more recently John Dewey, made interest the cornerstone of their philosophies of education. But psychologists did not begin studying interest until about the time of World War I; since then a great deal of significant research has been done in the measurement, development, and role of interest. We now know a great deal about its nature, and how to apply that knowledge to problems of human adjustment.

NATURE OF INTEREST

DEFINITIONS

In view of the confusion which has so frequently characterized discussions of interest, it will be well to devote attention to the nature of interest. There have been four major interpretations of the term, arising from four methods of obtaining data, which are here referred to as *expressed, manifest, tested,* and *inventoried* interest.

EXPRESSED INTEREST. An individual expresses his interest in a given object, activity, task, job, or situation verbally by saying that he likes, dislikes, or is indifferent to it. These verbal expressions are considered as identifying and delimiting the areas which interest the individual. Fryer (1931) reviewed the early work utilizing this definition of interest and concluded that the expressed or specific interests of children and adolescents are unstable and disappointing as diagnostic or prognostic data. The work of Lehman and Witty (1929, 1931, 1932) on the play interests of children confirms this conclusion, but does not warrant the sweeping condemnation of the use of interest inventories which confusion of terms led them to issue. Strong (1943) summarized studies of the constancy of responses to the four hundred items in his inventory over periods of from one day to 10 years, showing that the percentage of identical responses ranges from 52.6 for high school juniors after 6 years to 82.8 for women physicians after one day. The former finding agrees with Fryer's conclusions concerning children and adolescents, the latter shows that even specific expressed interests are rather stable in adults over a short period. The importance which may be attached to expressions of specific interests then varies with the maturity of the respondent and the purpose for which the data are to be used; the phrasing of the question must also be considered.

Manifest Interest. In view of the somewhat unreliable nature of expressions of interest, objective manifestations are frequently relied upon, and the term interest is used synonymously with activity. The fact that a young man developed photographs during his leisure time may be taken as evidence of interest in photography. Such manifest interests, however, have the disadvantage of sometimes being the result of interest in the concomitants or outcomes of the activity rather than in the activity itself: excelling in photography may have really been a means of achieving prestige in the eyes of friends. Then also, opportunities for manifestation of an interest may be limited, so that a genuine expressed interest may have no manifest counterpart. For these reasons, manifest interest has not been used as a predictor of interest in many investigations, although it has, oddly enough, been a commonly used criterion.

Tested Interest. If interest in an activity or vocation is likely to be manifested in action, whether in the election of a course in physics or in the reading of the story of the discovery of radium, it is also measurable through information retained on the subject in question. The amount and quality of information retained about a variety of subjects can be used as an index of degree of interest in each of these activities. During World War I, a few attempts were made to measure manifest interests (Fryer, 1931), but were not followed up as such because of the cumbersomeness of memory and information tests. Shortly before World War II work with a general information test was resumed by Flanagan (1939) and with a memory or attention test by Super and his students (1944). These beginnings led to the development of a general information test, differentially scored for various aircrew assignments, which proved to be the most valid single test for predicting success in Army flying training (Du Bois, 1947). Civilian applications of this technique are now being made.

Inventoried Interest. The relative unreliability of expressions of interest led to a search for trends underlying expressions of specific interests which might be more reliable than the single expressions themselves. Inventories, such as Strong's Vocational Interest Blank and Kuder's Preference Record, were the result. In these inventories expressions of specific interests are statistically combined to provide an over-all score indicating, in the case of Strong's Blank, the similarity of the respondent's basic interests to those of men who have been successful in various occupations, and, in the case of Kuder's, the relative strength of interest in nine different types of activity. Checking up on the effect of item instability on test scores, Strong (1943) found that shifts in response to as many as 125 of his 400 items had no appreciable effect on the total occupational score, the reason being that certain changes in expressed interest balanced other changes, leaving the underlying pattern or basic interest unchanged. This pioneer work of Strong's changed the whole pattern of research in the psychology of interest, made possible a great variety of studies, and led to the development of practical instruments for use in diagnosis, prediction, and counseling. Strong has brought together most of the important work with his inventory in a monumental volume (1943), and Super has summarized work with Kuder's inventory in a recent article in the Journal of Consulting Psychology (1947).

Applications of the Study of Interest. The applications of the study of interest are varied, and have affected such diverse fields as vocational guidance, personnel selection, the use of leisure, and the planning of school and college offerings. For example, Charters (1933) made a study of the postcollege activities of graduates of Stephens College in order to ascertain what should be taught in college to prepare students for adult life. The frequency with which various types of books are chosen by children is used as a guide in future selections for purchase or for class use. Studies to be referred to subsequently have led to the classification of occupations on the basis of interest and thus provided data essential to counseling and placement.

TYPES OF INTEREST

The isolation of basic or underlying interests by statistical methods has made possible not only the study of differences related to occupations, but also the study of differences in interest related to age, sex, nationality, and other factors. An understanding of the nature and extent of these differences is essential to the use of interest data.

Occupational Differences. The fundamental finding of Strong's early work in interest measurement was that the interests of men in a given occupation, say engineering, were different from those of men in general. This made possible the development of scoring keys for a variety of occupations whose workers were characterized by their own peculiar interests. Strong (1943) found that a few occupations are made up of persons with interests which do not differ from those of men in general, e.g., executives, but these occupations draw from a variety of distinct fields: an executive may be a sales manager, an engineer, or an accountant. Further work showed that certain occupational scales were closely correlated with each other, but uncorrelated or even negatively correlated with others. Strong therefore grouped the occupations in question in families based upon the intercorrelations between interest scales. These groups have been characterized as the biological science occupations (e.g., physician), physical science occupations (e.g., chemist), technical occupations (e.g., printer), social welfare occupations (e.g., Y.M.C.A. secretary), business detail occupations (e.g., accountant), business contact occupations (e.g., life insurance salesman), and literary-legal occupations (e.g., lawyer). Strong is not willing to name the first and third groups (1943) because the intercorrelations of the scales in those families are lower than those for other families, and suggests that as more occupational scales are developed these two groups may subdivide into several more homogeneous groups.

Sex Differences. Men tend to be more interested in physical activity, mechanical and scientific activities, politics, and selling than are women. Women tend to be more interested in music and art, literary activities, people, clerical activities, teaching, social work, and personal attractiveness (Strong, 1943; Lehman and Witty, 1929). Masculinity and femininity are scaled traits rather than dichotomies: some men are rather feminine, and some women are rather masculine, in their interests. Since interests are at least to some extent culturally developed, this may be a reflection of the fact that women are the carriers of culture rather than a proof that men who like art, literature, or people are effeminate: they may have been more exposed to culture and to human values than most men.

Age Differences. The interests of adolescents resemble those of adults, but tend to be somewhat less clearly defined, less humanistic, and more technical. As some fifteen-year-olds grow older, increasing maturity brings more interest in social welfare and less in engineering and related activities. By the end of adolescence interests are fairly well crystallized, and from age 25 to 55 there is relatively little change other than a decreasing interest in physical activity. These age differences made possible construction of a scale for appraising interest-maturity. Strong found that even several years of experience in occupations in which they were not interested did not change the interests of college graduates who were tested as students and retested as alumni (Strong, 1943).

Socio-Economic Differences. Men employed in higher level occupations are characterized by more literary and legal interest and less social welfare interest than are men at lower occupational levels. Men who have managerial responsibility tend to make high scores on Strong's occupational level scale, but so do men in certain other occupations not involving administrative work (Strong, 1943). Perhaps the occupational level scale is actually measuring level of aspiration. Research in the implications of socio-economic differences in interests has hardly been begun, and should prove fruitful.

Universality of Interest. In view of the extent to which occupational, sex, age, and socio-economic groups are differentiated by their interests, it is noteworthy that people's interests are in reality more similar than different (Strong, 1943), a fact which suggests that there may yet be hope for political and economic harmony. The correlation between the likes of fifteen-year-old boys and fifty-five-year-old men is 0.73, of college men and women 0.74, unskilled men and business and professional men earning more than $2,500 per annum 0.84; the interests of boys and women are least similar, and even this correlation is 0.48. Underneath the very real differences in interests which have been described, there is an even larger core of common interests.

PERMANENCE OF INTEREST

Specific expressed interest having been demonstrated to be somewhat ephemeral, considerable attention has been focused on the permanence of inventoried interest (Strong, 1943; Lehman and Witty, 1931).

As summarized by Strong, the research shows that inventoried interests have a retest reliability of 0.62 after one year in the case of eleven-year-old boys, 0.71 for seventeen- and eighteen-year-olds, and 0.80 for college freshmen. These data agree with those on change of interest with age, revealing greater stability of interest with increasing maturity. This is even more clearly brought out by Strong's follow-up studies of college students, the retest reliability for college *freshmen* after 9 years being 0.56, as contrasted with 0.71 for college *seniors* after 10 years. Furthermore, training and experience seem to have little effect on inventoried interests, for interest patterns found in adolescents by both Carter and Strong became more clearly defined as time went on, and even employment after college in a field in which low scores had been made did not result in an increase in those scores 9 or 10 years later.

Origin and Development of Interest. This independence of specific experience raises the question of the origin and development of interest. If it exists before vocational experience and is not affected by it, whence does it arise and why does it manifest itself as it does? Carter (1932) found a correlation of 0.50 between the inventoried interests of monozygotic twins, in contrast with one of 0.28 for dizygotic twins; Strong (1943) reports an average correlation of 0.29 between the inventoried interests of fathers and sons. These similarities might be the result of inherited aptitudes rather than of inherited interests. Indeed, Strong's theory is that the explanation of why vocational interests crystallize before occupational training or experience lies in the role which aptitudes play in their development. *Aptitude*, which results in success, produces satisfaction in the activities in which success is met. This satisfied feeling results, in turn, in interest in the activity. Since the same aptitude or combination of aptitudes causes success in related activities, a basic or underlying interest develops.

Personality has been stressed by other investigators as the source of interests. Darley (1938) is the principal exponent of this point of view, as a result of studies in which he found that social welfare interests tended to be related to feelings of inferiority, poor home and social adjustment, and lack of social skill, and that men with other occupational interests tended also to have certain personality traits in common.

A synthesis of both points of view has been achieved by Carter (1940) which seems to be acceptable to Strong (1943). Carter focuses on the adjustment of the individual to his environment and on the development and maintenance of self-concepts. The individual derives satisfaction from identification with a group, thereby attaining status. If his abilities permit, the identification is strengthened; if they do not, he must change self-concept, identifying with another group and developing a self-concept compatible with his aptitudes. Personality and aptitude go hand-in-hand to determine the selection of activities and goals and the consequent development of interests. Since personality and aptitudes begin to affect choices and outcomes in early childhood and have important differential effects on education and social life in adolescence, the reason for the crystallization of interest in adolescence becomes clear. By the middle teens the average boy or girl has developed attitudes toward his parents and toward others in authority which are likely to last a lifetime; attitudes toward his own contemporaries have begun to take lasting shape; success and failure in school subjects, leisure-time activities, and social life have been experienced and patterns of liking and disliking have been formed by satisfying and frustrating experiences. Preoccupational interests have begun to crystallize and, as adulthood is reached, these interests take on vocational form. Occupational training and experience therefore tend merely to confirm trends which have already been clear for some time.

MEASUREMENT OF INTEREST

The most productive approach to interest measurement so far has been that of the interest inventory. Begun at the Carnegie Institute of Technology after World War I, it was perfected by Strong at Stanford University. Tests of interests have had a more limited use and will therefore be treated more briefly.

INTEREST INVENTORIES

Although Strong, Kuder, Garretson, Symonds, Cleeton, and others have published vocational interest inventories, only a half-dozen authors have attempted adequately to standardize and validate their instruments; and only Strong and Kuder have made original contributions to the technique of the inventory. Similar to the vocational interest inventories is the Dunlap Academic Preference Blank designed for use with children in grades 6 through 9 for predicting academic success. This blank was subjected to an unusually rigorous and extensive validation over a period of 7 years. At this time, however, the Blank has not had sufficient use in the field to make possible a sound judgment as to its ultimate value. Only the Strong and Kuder inventories have been studied in a sufficient variety of situations over a long period of years to justify extensive comments on their use in counseling or selection.

STRONG'S VOCATIONAL INTEREST BLANK (Palo Alto: Stanford University Press). This inventory was revised in 1938. It now consists of 400 items, to each of which the respondent indicates the answer which best describes him. Occupations, school subjects, leisure activities, types of people are responded to in terms of liking, indifference, or disliking; self-descriptive traits are answered in terms of appropriateness or inappropriateness (yes or no).

The answer to each item is assigned a weight indicative of its significance for a given occupation. This weight is determined by giving the inventory to about four hundred men successfully engaged in an occupation, for example, Fellows of the American Psychological Association, and ascertaining what percentage indicate liking, indifference, and dislike for each item in the inventory. These percentages are then compared with the percentages of men-in-general who have been found to answer each item in each way, and a weight is assigned which varies with the difference between the two percentages. A man's score on the occupation, psychologist, is then the algebraic sum of the weights assigned to each of his answers by the psychologist scale, or, in other words, a picture of the degree to which his likes and dislikes differ from those of men-in-general in the same way as those of successful psychologists.

Strong's Blank can be scored for 37 occupations and three personality traits. For the former, numerical scores are converted into letter ratings ranging from *A* (interests like those of men in the occupation in question) to *C* (interests quite unlike those of men in that occupation).

The inventory was designed for use with college students and adults, but has been found useful in the last year or two of high school. It can even be used with fourteen-year-olds, as Carter (1941) has shown, but special caution is necessary to ensure that terms such as physiology and sociology are understood; and scores must be interpreted in the light of thorough understanding of the changes which are likely to take place during adolescence.

Because of the self-rating feature of inventories, it is possible for respondents to shift their scores in desired directions (Strong, 1943). Despite this fact, however, inventories such as Strong's have proved valuable not only in guidance but in selection, largely because the vocational significance of enough of the items is obscure to the average examinee.

The *validity* of Strong's Blank has been investigated by relating scores to those of other tests, to grades in school and college, to completion of training, to earnings and ratings of success, to persistence in an occupation, and to satisfaction in one's work. Because it has been the outstanding interest measure for about 20 years and because its author has maintained an active interest in his inventory, an unusual amount of validation material has been accumulated concerning it. All but the most recent studies are summarized in Strong's monograph (1943): for this reason, only a few of the major conclusions concerning the inventory's validity will be cited here.

In general, interest is unrelated to *mental ability*. Exceptions are scientific interest, with which it is positively correlated, and sales interest, the relationship with which is negative (Strong, 1943). Certain interests are related to personality traits, interest in sales work being related to social dominance, emotional adjustment, and desire for material welfare, whereas interest in social welfare work is related to poor social and emotional adjustment (Darley, 1938). Scientific and technical interests are somewhat related to mechanical aptitude (Strong, 1943).

Interest as measured by Strong's Blank has generally been found not to be correlated with *grades* in school or college (Strong, 1943). This appears to be due to the fact that when colleges select students partly on the basis of ability, and students select courses partly on the basis of interest, the quality of their work is relatively independent of degree of interest. Although a particular course may not appeal, the student can make himself do the work satisfactorily. Interest does play a part, however, in determining *continuation* in a field: student dentists who lack dental interests get grades as good as those of students with interests like those of dentists, but they have been found to be less likely to complete training (Strong, 1943). In one study, grades and interest were found to be positively related (Strong, 1944): in this instance, the students were selected by the Army on the basis of intelligence, and sent to training as personnel psychologists even though some had no interest in that subject. The correlation between scores on Strong's psychologist scale and grades in psychology was then as high as that between an intelligence test and grades in the same subject.

Success in selling life insurance was correlated with scores on Strong's life insurance salesman scale (1943), with the finding that 56 per cent of the men who rated *A* sold $150,000 worth of insurance per annum (yielding a $2,000 income to the salesman) while only 6 per cent of those who rated *C* wrote that much insurance. Production and interest in life insurance sales had a correlation of about 0.40. A somewhat similar trend was found for advertising executives, but cannot be said to hold for engineers, teachers, psychologists, and certain other groups for which Strong summarizes the findings (1943). This suggests that inventoried interest affects achievement in an occupation when the congeniality of the task is crucial to success in it, but not when its congeniality is less important than aptitudes, training, or other factors. A life insurance salesman must work at life insurance even during his leisure hours, making contacts and cultivating them; he must be systematic and not delude himself into thinking that nonproductive activity is productive work. A clerical worker, on the other hand, can go through the paces even though he has no special interest in his work, provided only that he has not an extreme distaste for it. It seems quite logical, then, that inventoried interest predicts success in certain occupations but not in others.

It has been pointed out by Darley (1941) that one of the outcomes of *persistence* in an occupation in which one lacks interest may be the development of compensatory hobbies. This hypothesis has been confirmed in the *Journal of Consulting Psychology* by Super (1940), who found, for example, that some model railroad engineers were men in nontechnical occupations who had technical interests. Stanford students followed up 5 and 10 years after graduation (Strong, 1943) were found, not only to choose occupations appropriate to their inventoried interests more often than not, but also to tend to shift from occupations in which they had made low scores to other fields in which they had made higher scores. Strong interprets this as evidence that interest is a cause of satisfaction or dissatisfaction in one's work. Combined with Super's findings, it suggests that there is something dynamic about interest, causing one to find an outlet for a dominant interest whether in the choice of a vocation, a change of vocation, or an avocation. And this, surely, is what interest should do rather than have a direct and noticeable effect on efficiency. The conclusion has received tentative confirmation in a small but significant study (Sarbin and Anderson, 1942) which showed that adults engaged in occupations appropriate to their inventoried interests were more likely to be satisfied than those otherwise engaged.

In using Strong's Blank for counseling, care must be taken not to put too much emphasis on a single high rating; rather, one should examine the ratings in related fields to see if they substantiate the one under consideration. Quite different interpretations should be made of ratings of *A* in engineering for two individuals, if one has ratings of *B*— and *C* for chemist and mathematician, and the second has ratings of *B*+ and *A* for the same occupations. This suggests examining an individual's scores for patterns of interest, for which purpose scores are grouped according to fields on the psychometric report. This is especially desirable as an aid in avoiding overemphasis on the few specific occupations scored in each family or group. An individual's primary patterns as defined by Darley (1941) consist of those fields in which he rates *A* or *B*+; his secondary patterns those in which he rates largely *B*+ and *B*; his tertiary patterns those in which he rates *B*— or less.

In interpreting inventory scores it is not enough to consider only the inventoried interest patterns. Equal consideration must be given to expressed interests (vocational choice and ambitions), manifest interests (previous employment, leisure-time activities), aptitudes, personality traits, and economic status. If ambition and primary interest pattern agree and if adequate achievement in related activities has been demonstrated, considerable confidence may be placed in the primary interest pattern. If, however, academic success has been limited while intellectual ability is adequate, it should be pointed out that the counselee has both the ability and the interest needed in the chosen field, but that he may have difficulty because of his inferior preparation. If both ability and achievement are inadequate for the preferred occupation, then related fields in the primary and secondary interest groups should be explored.

Evidence from the secondary pattern should be adduced whenever possible to support the primary pattern. It happens not infrequently that an individual's occupational choice does not agree with either his primary pattern or his abilities. His choice may be due to an induced interest resulting from parental pressure, financial reward, or desire for prestige, rather than from interest in the work itself. The evidence of inventoried interests and tested abilities can be brought to the consideration of the counselee, and he can be encouraged to explore related attitudes in counseling interviews and basic interests in further study or work.

If an individual's occupational choice falls in a tertiary pattern of interest, as it occasionally does, but his abilities relevant to the preferred field are great, he should be made aware of the apparent discrepancy and helped to consider its implications. If the preferred occupation is not one in which congeniality is essential to success he may do well despite divergent interests; these can find avocational outlets. Further exploration is desirable, however, and may lead him to change his plans. His primary and secondary patterns should suggest fields for exploration.

THE KUDER PREFERENCE RECORD (Chicago: Science Research Associates). The Kuder Preference Record was published in 1939 after Kuder had carried on related research over a period of 6 years. The inventory consists of 504 items, each of them a brief description of a certain type of activity. The activities are rather clearly vocational in nature; in general, they are more easily and more fully understood by high school students than are some of Strong's items, but for this very reason the inventory is probably more transparent than Strong's. Items are arranged in groups of three, each of a different type, so that in one trial one may have to indicate relative preference for a clerical, a musical, and a persuasive activity, and in another trial one's ranking of three scientific, literary, and clerical activities. In scoring the inventory, the sum of the preferences related to each of nine interest categories, or factors, is obtained, yielding scores for nine types of interest.

Unlike the procedure used in Strong's Blank, the method of constructing scoring keys used internal consistency as a criterion rather than relationship to an external criterion (Kuder, 1946; Super, 1947). For example, items which seemed to be literary in nature were examined to see if they tended to be liked by the same people. Those which were consistently answered were called the literary scale. Other items were examined in similar ways, to provide nine internally consistent and mutually independent scoring keys (mechanical, computational, scientific, persuasive, artistic, literary, musical, social service, and clerical). To many users of standard tests this procedure made the Kuder test seem subjective and lacking in validity. This was not actually the case, however, for this approach does demonstrate that the instrument measures nine relatively independent types of interest. To know what these interests really are, however, one needs external evidence of validity, occupational norms, for example. This has been somewhat slow in forthcoming, as in the case of most tests, but, as will be seen below, is available and adequate for certain purposes. The technique of scoring the inventory gives it a flexibility which facilitates validation research; it can, therefore, be expected that the usefulness of the Kuder will increase rapidly now that the preliminary work is completed.

The Kuder Preference Record is planned for use at the high school level and above. Although we do not know the effects of age on Kuder's scores, the instrument seems to be applicable to adolescents and adults; high school boys and girls enjoy taking it, and report that the items are meaningful to them. Adults generally consider the items appropriate. It seems probable that research will in due course show that the age changes which Strong's inventory has revealed will also be revealed with the Kuder inventory. Data on the effects of training on Kuder's scores are also lacking as yet.

The external *validity* of the Kuder inventory has been ascertained by methods somewhat similar to those used in studying Strong's, with the exception of follow-up studies. Considerable attention has been devoted to correlating Kuder's scores with Strong's scores, for one of the first questions asked was how well the instrument agreed with the then better known inventory (Triggs, 1944). Although there is a tendency for the keys of the two inventories to be related in expected ways, the relationship is not found in the case of all scales and is not perfect. To interpret Kuder's scores on the basis of what we know about Strong's Blank would therefore be unwarranted. The scientific and chemist keys correlate about 0.73, but the computational and accountant keys have correlations of only 0.48 and the literary and author keys only 0.28. The Kuder is not a cheaper substitute for the Strong, and must stand or fall on its own validity coefficients.

The relationship between Kuder's scores and school and college *grades* is about what it has been found to be for Strong's Blank. Grades are appropriately related to scores in certain cases but not in others, depending usually upon the range of interest in the group in question. Those who continue in a field of study tend to have higher scores in that field than do those who drop out. These and other validity studies have been summarized by Kuder (1946) and by Super in the *Journal of Consulting Psychology* (1947).

No studies have been published to date which relate Kuder's scores to *success* in occupations, but *occupational differentiation* on the basis of the Kuder inventory has been studied among 29 women's and 44 men's occupational groups. These groups, unlike Strong's, are in most cases very small and ill defined, but they provide a basis for judging the occupational significance of the nine interest scales. The patterns of interest scores reported by Kuder (1946) tend to differ according to expectation: accountants, for example, tend to make high scores in computational, literary, and clerical interests; and social workers score high in social service and literary interests. Kuder has begun to quantify these profiles to make possible occupational scores or indices; when this process has been carried further his inventory, like Strong's, will be useful not only in revealing a person's dominant interests but also in showing how his interests compare to those of successful persons in various occupations. As Kuder's instrument seems to measure women's interests better than Strong's (Super, 1947), this will be an important fact.

It has already been pointed out that there are no data which show the significance of Kuder's scores in subsequent vocational choice. Hahn and Williams (Kuder, 1946) have, however, produced evidence that women who are employed in work appropriate to their Kuder scores are better satisfied than those whose work was not compatible with their inventoried interests.

In *counseling* on the basis of the Kuder Preference Record, it is probably wise to keep in mind the relatively obvious significance of the items and the therefore relatively easily distorted scores. Special care should be taken in checking inventoried against manifest interests, by means of interviews and questionnaires. The lists of occupations relevant to the various types and combinations of interests provided by the manual should be used only as a tentative guide for exploration and discussion, with due recognition of the fact that the classification of occupations is still partly hypothetical. The user should both draw on a thorough knowledge of the psychology of interest as developed largely by Strong's work, and be cautious in making interpretations too far removed from the relatively limited validation research so far conducted with Kuder's inventory. Despite these limitations the Kuder Preference Record is a promising instrument, the value of which will no doubt increase as more research concerning it is published.

THE DUNLAP ACADEMIC PREFERENCE BLANK (Yonkers: World Book Co.). This inventory was released for general use in 1939. There are two forms of the blank each containing ninety items, which may be scored differentially for the following areas: literature, geography, word meaning, paragraph meaning, language usage, arithmetic, history, general ability, and intellectual alertness.

The basic assumption underlying these blanks is that expressed preferences for certain words or phrases in one's academic environment will reveal something of one's achievement in various academic subjects. Validational studies against standard achievement tests support this hypothesis, since the validity coefficients ranged from 0.28 to 0.72 for the various areas with a median value of 0.52. The median test-retest reliabilities were 0.82 for Form A and 0.90 for Form B. That the Blanks are not disguised intelligence tests was shown by a study in which individuals were matched as to intelligence but differed greatly as to achievement. It was found that these differences in achievement corresponded to differences in the preference scores (Sharkey and Dunlap, 1940).

INTEREST TESTS

Although interest tests were first tried out during and shortly after World War I in the form of information tests, work with the technique was virtually abandoned until about 1940. At that time the Cooperative General Culture and Contemporary Affairs Tests (New York: Cooperative Test Service) and the Michigan Vocabulary Profile Test (Yonkers: World Book Co.) were published. Work with a General Information Test in the Army Air Forces (Du Bois, 1947), including information about flying, subsequently demonstrated the practicality and validity of the technique. Similarly, in 1941 Dunlap and Wantman prepared a "Desire to Fly Inventory" based on information about aviation for the Committee on the Selection and Training of Aircraft Pilots. The bi-serial correlation between scores on the Inventory and passing or failing in primary flight training was 0.35.

The underlying assumption of the information test of interest is that a person will attend to and retain that which interests him. True, he will notice and learn certain things because circumstances force him to do so, but it is assumed that in a large enough sampling of miscellaneously acquired information these noninterest factors will cancel out and differences in types of information retained will reflect differences in interest. Thus Greene's Michigan Vocabulary Profile Test presupposes that someone interested in medicine will learn more about recent developments in science than will someone not so interested. The AAF General Information Test assumed that men interested in flying would have, among other interests, basic athletic interests which would cause them to learn more about certain types of sports than do men not interested in flying. In the AAF program testing was done before admission to flying training; the validity of the assumptions underlying the test is suggested by its correlation of 0.50 with success in flying training (Du Bois, 1947). The hypothesis may, it is true, have been mistaken, but the test items written with it as a basis were remarkably successful predictors of flying success. Adaptations of this technique have already been made for use in industrial personnel work; more will no doubt be made, and in due course, instruments should be available which are more adequately normed and validated for counseling than the pioneer tests now in use.

INTEREST AND ADJUSTMENT

We have seen that early research in the psychology of interest was directed toward the improvement of predictions of success in the hope that tests of interest and aptitude would so supplement each other as to make possible accurate predictions of educational or vocational performance. When this hope proved generally unjustified, attention was shifted to the relationship between interest and continuation in an activity. In this instance the hypothesis was confirmed, and we have seen that students are more likely to graduate from professional schools and men are more likely to remain in an occupation if their interests are like those of others who have been successful in the field in question. Finally, studies of job satisfaction and of the relationship between vocations and avocations have demonstrated the dynamic quality of interest and its fundamental role in human adjustment. As measuring instruments have been perfected and psychologists have become more sophisticated in their use, it has been found that interest can affect not only satisfaction but even achievement in an occupation. In vocational counseling and selection there will continue to be an emphasis on predicting success, in which the insightful use of improved inventories and tests will therefore apparently play an increasingly important part. But the attention to general adjustment, morale, and satisfaction which characterizes the best vocational counseling and personnel work of today will, to an even greater degree, result in widespread study of the psychology of interest and extensive use of the instruments with which interest is measured.

READING REFERENCES

Carter, H. D., *Vocational Interests and Job Orientation.* Palo Alto: Stanford University Press, 1944.

Darley, J. G., *Clinical Aspects and Interpretation of the Strong Vocational Interest Blank.* New York: The Psychological Corporation, 1941.

Strong, E. K., *Vocational Interests of Men and Women.* Palo Alto: Stanford University Press, 1943.

Super, D. E., "The Kuder Preference Record in Vocational Diagnosis." *Journal of Consulting Psychology,* 1947, **11,** 184–193.

———, "Vocational Interest and Vocational Choice." *Educational and Psychological Measurement,* 1947, **7,** 375–384.

17. APTITUDE AND INTELLIGENCE

GEORGE K. BENNETT, PH.D.

President, The Psychological Corporation; Vice Chairman, Division of Anthropology and Psychology, National Research Council; Member, American Board of Examiners in Professional Psychology; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology. Formerly: Lecturer in Psychology, Columbia University; President, New York State Psychological Association; President, Division of Industrial and Business Psychology of the American Psychological Association.

Intelligence has been defined in various ways by those who have concerned themselves with this problem. One authority, in a facetious mood, said that intelligence was that characteristic which was measured by intelligence tests. Other psychologists have described it as the capacity to learn or the ability to do abstract thinking. An enterprising student might find and classify many such definitions, possibly hundreds. There is, however, a common core of meaning in the various definitions. Intelligence may be regarded as the ability of an organism to acquire, from its own experience, modes of behaving which in general operate to enhance the likelihood of its survival.

Intelligence is less difficult to define from an operational standpoint. An intelligent rat will learn where to get food in a maze more rapidly than a less intelligent rat. Given comparable environment a child of superior intelligence will have a larger vocabulary at a given age than will one with less of this characteristic. A bright high school graduate will be more likely to succeed in college or technical school than a duller classmate.

There is no very clear distinction between intelligence and aptitudes. In fact, psychologists often refer to intelligence as "academic aptitude" because of its close relation to school success. Aptitude refers to potentiality for future ability, usually after training, and for a specific type of activity. Thus, a boy in high school may be found to have aptitude for mechanical work although he has not been trained for it.

Prediction is the goal of intelligence and aptitude tests. When a baby being considered for adoption is given an intelligence test, the purpose is to estimate how well the child will be able to learn and think so that he may be placed in the sort of home that will offer him appropriate opportunities. When a newly inducted soldier is given the Army classification test, the score is used in the prediction of his ability to comprehend and carry out instructions. Experience records, such as high school grades or the reference of a former employer, may also be used as predictors of some types of performance but these data are often difficult or impossible to obtain and usually hard to interpret.

A number of very capable investigators have tried to discover the basic units of mental ability. In 1904 Charles Spearman of the University of London announced the "Theory of Two Factors." Calling attention to the tendency of test scores to be positively correlated, he proceeded with a mathematical analysis leading to the postulation of one general factor of great importance and many specific factors. Every intellectual activity involved the operation of the general factor plus one or several specific factors. Different activities would utilize different specific factors but always the same general factor. The "Theory of Two Factors" aroused some controversy and stimulated further experimentation. The theory was criticized by Godfrey Thomson on logical and mathematical grounds, while it was supported by others. Among those responsible for important research in this field are K. J. Holzinger, T. L. Kelley, Harold Hotelling, and L. L. Thurstone. More recent mathematical developments have permitted "factorial analysis" which attempts to isolate and identify the elements of human ability. Thurstone, for example, has identified nine "primary abilities" including visual or spatial, perceptual, numerical, logical or verbal relations, verbal fluency, memory, inductive ability, restricted problem solution, and deductive ability. Unfortunately, these factors are not sufficiently stable; nor are they comprehensive enough for general use. If more refined methods of analysis establish dependable factors, an enormous advance will have occurred in this field of psychology. At this time, the practical measurement of intelligence and aptitude proceeds with little dependence upon these theoretical investigations.

Whereas factorial analysis attempts to isolate independent elements of intellectual behavior, some of which are difficult to define or interpret, a battery of directly meaningful tests, not wholly uncorrelated, seems to offer a satisfactory means for appraising the relative strengths and weaknesses of the individual in different areas of endeavor. When a sufficient body of validational data has been obtained, the relative merits of different approaches can be more objectively appraised.

THE MEASUREMENT OF ABILITIES

INTELLIGENCE

The measurement of intelligence is usually regarded as having begun with the work of Alfred Binet, who in 1896 developed a test series and in 1908 published the basic elements of his scale for measuring intelligence. Credit should also be given to a number of other investigators. Notable among these is J. McKeen Cattell, who established a series of tests for use with Columbia University students in 1891. While some of Cattell's tests, such as sensitivity to pain and reaction time to sound, are not now regarded as measures of intelligence, others still are; and he did not have the benefit of fifty years' perspective.

In 1916 the Stanford Revision of the Binet-Simon Tests appeared. Lewis M. Terman of Stanford University had adapted the Binet material for American use and prepared and standardized an individual intelligence scale suitable for both children and adults. The revision consisted of twelve levels, ranging from three years to superior adult. At each level the subject was confronted with several types of problems, usually six. According to the number of problems satisfactorily answered at the various levels, a score in terms of years and months of "mental age" was obtained. In the case of children the mental age was divided by the actual or "chronological age" to yield the "intelligence quotient" (usually designated as I.Q.). Since mental age and "I.Q." are concepts readily understood by those concerned with the education of children, the Stanford Revision rapidly gained adherents. It has been extensively used with normal and problem children and to a lesser extent with adults. It has been of great value in the diagnosis of school failure and in the detection of feeblemindedness. In 1937 Terman, with the assistance of Maud Merrill, prepared an improved version (Revised Stanford-Binet Scale) which retained many of the characteristics of the 1916 scale.

Other American revisions of Binet's work were prepared by H. H. Goddard, J. P. Herring, and F. Kuhlmann. None of these attained the popularity of Terman's scale.

When the United States entered the World War in 1917, a Divison of Psychology was established under the direction of the Surgeon General of the Army and headed by R. M. Yerkes. Most effort was directed toward the establishment of intelligence scales suitable for simultaneous administration to groups of subjects, since individual testing was not practicable for the numbers of men to be inducted. Two such scales were developed, Alpha for literates and Beta for the illiterate and foreign language groups. This work was so excellent that both scales are still in use and many of the more recently published group tests have much in common with Alpha. Many psychologists now eminent in the field of tests were associated with this effort.

The success of group tests in the Army led to their rapid spread, first in schools and to a lesser extent in industry. While these intelligence tests were suitable and valuable for student classification, their industrial application was frequently followed by disappointment. As experience accumulated, it became apparent that the abilities measured by these tests were not of primary importance in most jobs. The difficulty lay less in the tests themselves than in their inept and inappropriate use.

At the present time there is a great variety of individual and group intelligence scales, most of which have features recommending them for specific situations. Tests are available for infants of three months and for the aged. Others are suitable for use with the deaf, blind, and mentally disordered. Still others, such as the American Council on Education Psychological Examination by the Thurstones and Thorndike's CAVD, can be used with exceptionally superior individuals.

MECHANICAL APTITUDE

Tests of nonacademic abilities have a history almost as long. Among the earliest of these is the Mechanical Assembly Test designed by J. L. Stenquist. Originally used in the Army testing during World War I, it was subsequently released for general use.

This test requires the subject to reassemble ten common mechanical objects, such as a cupboard catch, a bicycle bell, and a mousetrap. It has been modified and extended by those working on the Minnesota Mechanical Ability project. The Stenquist Assembly Test and some trade tests were used to direct demobilized soldiers in World War I toward appropriate civilian activities. Paper-and-pencil tests of mechanical knowledge and ingenuity have since appeared. Among the most frequently used of these tests are those by Stenquist, O'Rourke, Bennett, MacQuarrie, and Likert and Quasha. Form-board tests intended to measure perception of spatial relations require the subject to place odd shaped blocks in their corresponding apertures. Others, measuring speed of movement, consist of a number of cylindrical blocks to be turned over or to be placed in round holes in predetermined order. Still others, designed for finer coordination, include the placing of brass pins in small openings by means of fingers or tweezers.

CLERICAL APTITUDE

Clerical aptitude tests resemble intelligence tests. This resemblance is natural since there is much in common between clerical tasks and school work. The academic skills of reading, computation, and writing are highly important in most clerical jobs. Superior command of the elementary processes is more advantageous than the more diverse culture of the advanced courses. A bookkeeper or statistical clerk will need proficiency in the fundamentals of arithmetic but will profit little from algebra and less from trigonometry. Clerical aptitude tests ordinarily contain several sections, each of which covers some specific type of clerical content, such as spelling, grammar, form completion, or classification. Examples of such tests are the Detroit Clerical Aptitudes Examination, the Psychological Corporation General Clerical Test, and the Thurstone Examination in Clerical Work. O'Rourke's Clerical Problems and Clerical Reasoning Tests represent another approach to the measurement of clerical ability. The first of these involves six types of simulated clerical activity, while the second is primarily an intelligence test with appropriate content. Name and number checking comprise the Minnesota Vocational Test for Clerical Workers. Speed and accuracy in these simple functions are reasonably good predictors of success in the majority of clerical jobs.

All clerical tests are partial tests of aptitude and partial measures of proficiency since office work is largely dependent on knowledge of school subjects. In many instances an achievement test of tenth grade level would produce equally good predictions. For some of the more specialized clerical skills the general tests are less satisfactory. Typewriting ability does not correlate well with over-all intellectual level, while verbal facility bears little relationship to comptometer operation. A number of fairly satisfactory tests of clerical proficiency exist for use with trained individuals. The Seashore-Bennett Stenographic Proficiency Tests consist of graded series of letters recorded on phonograph discs. The subject listens, takes stenographic notes, and subsequently transcribes each letter in typewritten form. Performance on this work-sample corresponds well with independent ratings of proficiency.

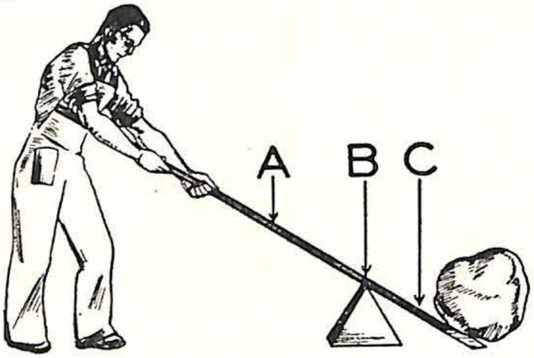
APTITUDES FOR THE PROFESSIONS

Aptitude for the professions has received much attention because of the interest of professional schools and associations in raising the caliber of students and members and in eliminating overcrowding. Most professional schools require candidates to have completed all or a major part of a liberal arts course before admission. Many scrutinize their applicants further for the possession of (1) ability to comprehend the content of the professional course, (2) interest in and facility with the types of problems to be encountered, and (3) ethical principles and industriousness consistent with professional standards. Aptitude tests are commonly used in schools of medicine and law and to a lesser extent in teachers colleges and schools of dentistry. The Association of American Medical Colleges has, since 1930, recommended to its members the use of a special medical aptitude test, a new form of which is prepared annually. This is primarily a difficult test of academic aptitude, with emphasis on scientific and biological aspects. The law schools have not unified their selective methods to so great an extent. A number use the Law Aptitude Examination by Ferson and Stoddard, a test requiring primarily logical reasoning with verbal symbols. Others have developed specific batteries of tests for their own applicants.

Until recently, special programs for the selection of engineering students were few. Most engineering schools seemed to rely upon secondary school grades with particular emphasis on marks in mathematics and science courses. The Graduate Record Examination Bureau recently has developed an engineering program to remedy this deficiency. The Yale Educa-

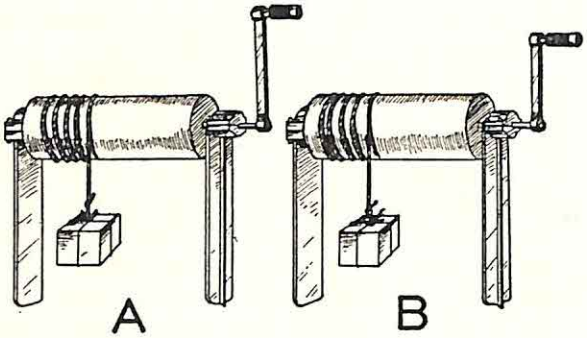
DRAW A CIRCLE AROUND THE RIGHT ANSWER.

1.



If a man were lifting a stone with this crowbar, at which point would the bar be most likely to break?

A B C



With which windlass can a man raise the heavier weight?

A B Equal

2. "When I say 'Go' mark out the right word in each line and be sure to put the number of the right answer in the parenthesis. Work hard! Ready, GO!"

( )	Weld means to	1 fuse together	2 bore	3 harden	4 melt
( )	A tool is tempered by	1 compression	2 heating	3 melting	4 welding
( )	A lock washer is used with	1 bolts	2 nails	3 rivets	4 screws
( )	A dowel is used to	1 dig with	2 smooth	3 fasten	4 bore

3. First look at Problem 1. There are two parts in the upper left-hand corner. Now look at the five figures labelled A, B, C, D, E. You are to decide which figure shows how these parts can fit together. Check this figure.

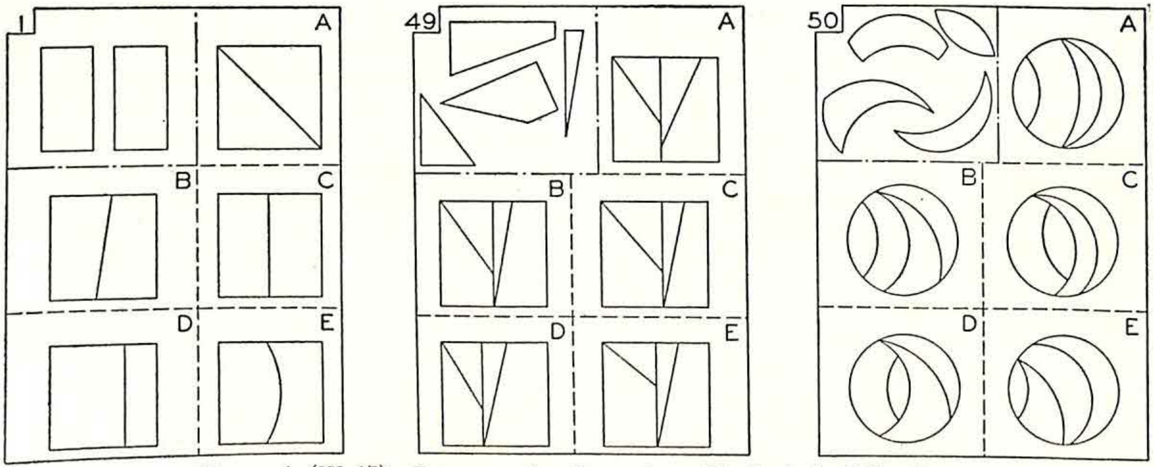


Figure 1 (III–17): *Representative Items from Mechanical Ability Tests*

(1) Mechanical Comprehension Test (The Psychological Corporation); (2) Detroit Mechanical Aptitudes Examination (Public School Publishing Company); (3) Revised Minnesota Paper Form Board (The Psychological Corporation)

NUMBER COMPARISON

√ Check if the two numbers are the same.

307......309
4605......4603
976......996
101267......101267
3065432......30965432
103345700......103345700
46754......466754
3367490......3367490
2779......2778
57394......57394

NAME COMPARISON

√ Check if the two names are the same.

Hulme Co.......Hulne Co.
L. T. Piver......L. T. Piver
Foley & Co.......Foley & Co.
Floyd Gibbons......Floyd Gibbens
Dole Publishing Co.......Dole Publishing Co.
Chase Bag Co.......Chase Bag Co.
Aladdin Co.......Aladdin & Co.
Arthur Bier & Co.......Arthur Bier Co.
Mydall Cain......Mydell Cain
M. C. D. Borden & Sons......M. C. D. Borden & Sons

SPELLING

Some of the words below are correctly spelled and some are not. Where a word is **not** spelled properly, write the correct spelling in the space following the word. Do nothing when a word is spelled correctly.

fourty..........................
vacant..........................
niece..........................
nuisence..........................
councel..........................
tenement..........................
apoligize..........................
remittance..........................
foreign..........................
elegible..........................
aquainted..........................
official..........................
inteligence..........................
presious..........................
performance..........................

PROBLEMS

Do these problems in any order you like. If you find one difficult, go on with others and come back to it later. All are quite straightforward; there are no "catches."

1. Any correspondence on lighting has to be sent either to Mr. Jones or to Mr. Smith. Any correspondence on heating has to be sent to Mr. Smith or to Mr. Robinson. A joint estimate for heating and lighting arrives, to whom should it be sent?

10. A clerk has still half an hour to work before the office closes. There are four jobs she could do—the filing would take her 12 minutes, typing letter A would take her 5 minutes, typing letter B would take 45 minutes, and typing a list 18 minutes. Underline the words describing the jobs which she should do before the office closes, in order to fill in her time.

12. The result of multiplying 5 by (   ) is the same as that of dividing 60 by (   ) or of adding 17 and 13. Write the missing figures.

Figure 2 (III–17): *Representative Items from Clerical Ability Tests*

Number and name comparisons from the Minnesota Vocational Test for Clerical Workers (The Psychological Corporation); spelling from the Stenographic Aptitude Text (The Psychological Corporation); and problems from the N. I. I. P. Clerical Test, American Revision (Herbert Moore).

tional Aptitude Testing program includes several tests of specific validity for engineering courses.

Specific tests have been developed in the fields of art and music. In general these have measured the sensitivity of the subject rather than his creative capacity. Among the better known art tests are the Meier (a revision of the Meier-Seashore), the McAdory, and the Lewerenz. A more recent test is the Graves Design Judgment Test which attempts to measure perception of abstract aesthetic principles. In music two sets of phonograph records, the Seashore and the Kwalwasser-Dykema, are used to determine the extent to which the individual possesses certain basic capacities for discrimination. There are several tests measuring different phases of musical ability. The difficulty of rating the products of performance in these fields has limited most of the tests to the appreciative aspects.

Many schools of nursing have found it profitable to utilize a candidate testing service established in 1933 by Edith M. Potts. This service examines scholastically eligible applicants and reports regarding their general ability and specific knowledge. Schools of nursing which have used this service for several years report an appreciable diminution in the proportion of academic failures. A similar service is also offered by a division of the National League of Nursing Education.

THE USE OF TESTS IN PREDICTING VOCATIONAL SUCCESS

From the standpoint of the school or the employer, the most important item of information regarding an applicant is his probability of success in the course or job contemplated. The applicant, however, should be interested in whether this educational or occupational opportunity is that in which he is likely to receive the greatest degree ot success and satisfaction. While present tests can answer the school or employer with some benefit in the majority of cases, it is revealing to consider the basis for choice on the part of the individual.

If the counselee is a high school graduate with a superior record, there are many courses open to him. He can go on to a liberal arts college or an engineering school, he can undertake vocational training for a variety of occupations, or he can apply for immediate employment. The judicious use of tests can be of some service since intelligence tests will indicate the desirability of higher education and if this appears possible, achievement tests in mathematics may help in steering him toward or away from such highly technical work as engineering. If college is not indicated, mechanical and clerical tests may narrow the vocational choices. However, if he is liberal arts college material, the chances are that he can ultimately meet the minimum requirements for admission to a variety of occupations or professional courses and, with the information now available, it is impossible to say definitely where his best chance for success will be. A competent guidance counselor will consider the other factors in the situation: including interests, family occupations, financial status, and occupational opportunities, and from these and the test data arrive at some reasonable and feasible plan. Such a plan usually provides for supplementary education or training in line with the counselee's strongest aptitudes, but the judgment of the adviser bears a major burden. This is due to the absence of prognostic tests which will, for example, discriminate clearly between clerical workers and salesmen, electricians and plumbers, or teachers and lawyers.

Differential aptitude tests, validated against a variety of scholastic and occupational criteria, have begun to appear since the end of World War II. Such batteries, consisting of several highly reliable and relatively independent measures standardized upon the same population, yield profiles of scores that reveal the idiosyncrasies of the individual in a manner which gives important assistance to the counselor. As research in this direction continues, the tools of the counselor will become more effective.

GUIDANCE AND SELF-GUIDANCE

When tests are used to help a person find the type of education or occupation for which he is best suited the process is called guidance. Here it is usually necessary to find the individual's scores on a number of tests, including one or more of intelligence, several of special aptitudes, plus others of achievement, interest, and personality. The guidance counselor also takes into consideration the pertinent elements of the client's past experience and present situation to outline a course of action. The final advice frequently depends upon the opportunities for training or employment. The competent psychologist will rarely recommend a vocation which is on the down grade or one which is so overstaffed that the competition for existing jobs is intense. On the other hand, a good adviser seeks to direct his clients toward fields in which the likelihood of satisfactory employment will be great.

Self-guidance is impracticable for most of those who might expect to benefit by it. The person in need of guidance can scarcely be expected to choose properly the tests pertinent to his problem. He is even less likely to possess the knowledge requisite for interpretation of the scores. Finally, he faces the almost insurmountable obstacle of attempting to maintain an objective attitude toward his personal problems. While the ideal of self-guidance has much to commend it, the difficulties of preparing persons to advise themselves seem to be of much the same order as teaching every man to prescribe for his own physical ailments.

Fortunately, the need for self-guidance will diminish as publicly supported agencies become more numerous. At present many colleges and secondary schools have expert services available to their students and such organizations as the Y.M.C.A., Y.W.C.A., Y.M.H.A., religious agencies, community service centers, private consultants and, in some states, public employment services render vocational counsel to their clients.

The need for proper classification and assignment of men in the Armed Forces will have great influence on the future of guidance. The services have trained many men as interviewers and classification officers. The new recruits have been exposed to the elements of vocational guidance, most of them for the first time. Industry and education in the future will be staffed to a large extent by men who know from experience that testing and guidance can aid in the choice of a career.

A major difficulty in the field of aptitude test construction is the lack of dependable criteria of performance in most occupations. In test construction it is essential that some later measures of proficiency be obtained so that the predictive value of each test element, or item of personal data obtained beforehand, can be empirically determined. In schools, grades provide reasonably satisfactory estimates of proficiency, but most occupational situations are devoid of adequate criteria. This lack is widely recognized among psychologists, and conscientious test constructors are constantly searching for better measures of proficiency against which tests may be evaluated.

While the present status of intelligence and aptitude testing falls far short of providing clear and specific indications for all who seek vocational guidance, it can supply important information in many situations. When a grammar school graduate must choose among general, commercial, and technical high school courses, judiciously selected tests will be of profit. When a high school graduate faces a similar dilemma, other tests will assist him. Apprentices selected by tests are superior to those chosen by other methods. As more research in this field is completed, we shall come closer to the goal of being able to provide the individual with a close estimate of his chances of success in any vocation.

READING REFERENCES

Bennett, G. K., and R. M. Cruikshank, *A Summary of Clerical Tests*. New York: The Psychological Corporation, 1949.

Bingham, W. V., *Aptitudes and Aptitude Testing*. New York: Harper & Brothers, 1937.

Cox, J. W., *Mechanical Aptitude: Its Existence, Nature and Measurement*. London: Methuen and Company, 1928.

Crawford, A. B., and P. S. Burnham, *Forecasting College Achievement: A Survey of Aptitude Tests for Higher Education:* Part I. *General Considerations in the Measurement of Academic Promise*. New Haven: Yale University Press, 1946.

Garrett, Henry E., and M. M. R. Schneck, *Psychological Tests, Methods and Results*. New York: Harper & Brothers, 1933.

Greene, E. B., *Measurements of Human Behavior*. New York: The Odyssey Press, Inc., 1941.

Hull, C. L., *Aptitude Testing*. Yonkers-on-Hudson: World Book Company, 1928.

Paterson, D. G.; G. Schneidler; and E. G. Williamson, *Student Guidance Techniques*. New York: McGraw-Hill Book Company, Inc., 1938.

———; R. M. Elliott, et al., *Minnesota Mechanical Ability Tests*. Minneapolis: University of Minnesota Press, 1930.

Pintner, R., *Intelligence Testing: Methods and Results* (rev. ed.). New York: Henry Holt and Company, Inc., 1931.

Spearman, C., *The Abilities of Man*. The Macmillan Company, 1927.

Terman, L. M., *The Measurement of Intelligence*. Boston: Houghton Mifflin Company, 1916.

Thomson, G. H., *The Factorial Analysis of Human Ability*. Boston: Houghton Mifflin Company, 1939.

18. PERSONALITY AND PHYSIQUE

OTTO KLINEBERG, M.D., PH.D.

Associate Professor of Psychology, Columbia University; Director of the Tensions Project and Acting Head of Department of Social Sciences, UNESCO. Formerly: Professor of Anthropology, Sarah Lawrence College; Visiting Professor of Psychology, University of São Paulo, Brazil; Chairman, Society for the Psychological Study of Social Issues; President, Eastern Psychological Association and Sociedade de Psicologia de São Paulo; Representative of American Psychological Association to National Research Council; Committee on Negro Studies, American Council of Learned Societies; Committee on African Anthropology, National Research Council; Principal Social Psychologist, Federal Communications Commission; Regional Specialist for Germany, OWI; Special Scientific Consultant, War Department.

PHYSIOGNOMY AND PHRENOLOGY

The attempt to explain variations in personality on the basis of differences in physique had its origin in the very earliest speculations concerning human behavior. At one time this took the form of an argument from analogy with the physical appearance of animals, a man with a leonine head having the courage of a lion, one who looked like an ox, having the characteristics of an ox, and so on. Aristotle, who took issue with this type of analogical reasoning, nevertheless believed strongly in the parallelism between physical and mental characteristics, and gave many examples of such a relationship.

The more recent history of this problem is associated with the name of Lavater (1783), who published an extensive series of sketches in which physiognomic features were used as guides for the understanding of personality. Lavater's views have been almost completely discarded by modern psychologists, although popular beliefs along similar lines continue to exercise an important influence and have even been used in a system of vocational guidance (Blackford and Newcomb, 1914). The scheme of interpretation established by phrenologists (Gall, 1810; Spurzheim, 1815; Fowler, 1873) has similarly disappeared from scientific writings.

The most recent objective attempts to relate personality to physique admit the impossibility of establishing any direct correspondence between personality characteristics and single physiognomic features, such as the size of the nose or the shape of the head, and look instead for dynamic interrelationships between the functioning of the organism from both the physical and behavioral standpoints.

ENDOCRINE GLANDS

The glands of internal secretion seemed for a time to furnish a satisfactory bridge between physique and mentality. The effects of the secretions appeared to manifest themselves simultaneously in the organism and in behavior, so that inferences could be drawn from one to the other. This constitutes the basis of the scheme erected by Berman (1921), who describes various types of personality in terms of the hyper- or hypofunctioning of a single gland.

The present position in the field of endocrinology is a more conservative one. It holds that whereas in pathological cases, such as cretinism or Addison's disease, the effect of endocrine disturbance upon behavior is unmistakable, such an effect cannot be demonstrated in the case of normal variations in glandular functioning. The attempts, for example, to relate variations in basal metabolism to differences in speed of motor or mental activity (Lanier and Leedy, 1933; Steinberg, 1934) have yielded conflicting results. One of the few studies which represent positive findings is that of Mateer (1934), who reported rapid improvement in the intelligence quotients of hypopituitary cases with proper training, teaching, and endocrine therapy. In this study, however, the subjects could not be regarded as normal.

Two important considerations must be kept in mind in any interpretation of the relation between endocrine secretions and behavior. The first is that endocrine functioning is not to be explained entirely in terms of heredity; it may be markedly influenced by factors in the physical environment as well as by the experiences of the individual. Emotional disturbances may, for example, influence the functioning of the thyroid or the adrenal, sometimes bringing about lasting changes in the strength of the secretions (Dunbar, 1938). Such environmental influences reduce the possibility of using endocrine variations as an explanation of differences in personality.

The second consideration is that the individual endocrine glands do not function in isolation, but together form an interdependent system. Sexual maturation, to take only one instance, appears to be dependent upon changes in the adrenal cortex, the thyroid, and the thymus, as well as in the sex glands.

CONSTITUTIONAL TYPES

Explanations of personality have also been made in terms of general bodily constitution. It is probable that Hippocrates was the first to use constitutional types as an aid to understanding the variations in individual susceptibility to different diseases, and his *habitus phthisicus* and *habitus apoplecticus* are used by certain medical clinicians as a partial explanation for the development of pulmonary tuberculosis and cardiovascular disorders, respectively.

The writings of Kretschmer (1931) and Sheldon and his collaborators (1940) represent the best known psychological contributions in this field. Kretschmer was concerned primarily with the relationship between constitutional type and mental disease. His results appear to demonstrate that the relatively short, thickset "pyknic" type shows the greatest susceptibility to manic-depressive psychosis, and the relatively tall "leptosome" or "asthenic," the well-developed, muscular "athletic," and the abnormal or "dysplastic" types tend to be found among patients suffering from schizophrenia or dementia praecox. Several investigations (Garvey, 1933; Burchard, 1936; Farber, 1938) have indicated, however, that this relationship is not so close as Kretschmer suggests, and that when patients are equated for age, most of the correspondence between physique and psychosis seems to disappear.

Kretschmer believes that the correspondence between his constitutional types, on the one hand, and psychological characteristics, on the other, can be demonstrated for normal persons, the pyknic type being found among those who may be described as cyclothyme, or extrovert, and the other physical types among those who tend to be schizothyme, or introvert. An attempt in this country to check his hypothesis failed, however, to yield any positive findings (Klineberg, Asch, and Block, 1934). In spite of the great interest which Kretschmer's theory has aroused, it does not seem at the present time to be a particularly useful instrument for the interpretation of variations in normal personality.

Sheldon (1940) presents an elaborate and carefully constructed scheme for the description of the varieties in human physique, as well as their relation to temperamental characteristics. He identifies three structural components of bodily constitution: (1) endomorphy, which means relative predominance of soft roundness throughout the various regions of the body; (2) mesomorphy, relative predominance of muscle, bone, and connective tissue; and (3) ectomorphy, or relative predominance of linearity and fragility. Numbers from one to seven are assigned to the relative positions of the individual with reference to each component. For example, a 711 is a physique extreme in endomorphy and at a minimum in the other two components; a 444 represents an individual who is at the midpoint on all three scales, and so on. These three components are closely related to three basic aspects of temperament: (1) endomorphy is related to viscerotonia, which is roughly identifiable with love of comfort, relaxation, sociability; (2) mesomorphy is related to somatotonia, characterized by the will to exertion, exercise, and vigorous self-expression; and (3) ectomorphy is related to cerebrotonia, which refers to the attentional and inhibitory aspects of temperament, and to the domination of the sensory and central nervous systems.

In connection with this treatment of physical types, Sheldon discusses a number of practical consequences of the interdependence of constitutional and temperamental components. He suggests that an understanding of the varieties of human physique is important, not only in connection with susceptibility to various forms of physical and mental disease, but also in the case of crime and delinquency, sexual attraction, differential education, preferences for certain types of dress or furniture, and many other aspects of individual and social behavior.

Some of the evidence for this interdependence is presented in a later volume (Sheldon and Stevens, 1942), in which the varieties of temperament are described with illustrative case studies. The material is exceedingly interesting, but as yet not entirely convincing. The conclusions are, however, so important that it would be desirable to have the study repeated by other investigators. Judgment as to the final value of Sheldon's scheme must be reserved for the time being.

PERSONALITY AND RACIAL TYPE

There is a popular impression to the effect that the inherited physical characteristics which are used in classifying racial groups also have certain implications for personality. Not only the striking physical differences between Negroes and whites, but even the minor variations which frequently distinguish, for example, Italians from Englishmen, have been interpreted by laymen and some scientists as having a direct relation to variations in intelligence and temperament.

The present consensus among scientists is that there has never been any really adequate demonstration of such a relationship, and that there is no scientific justification for assuming that racial type determines personality. There is just about the same distribution of inherited temperamental characteristics in any one racial or national group as in any other. As the result of training and tradition, certain types of behavior vary in frequency in different communities, but physical type has apparently little or nothing to do with the case (Klineberg, 1935).

PHYSIQUE AND CRIME

Among the alleged social and psychological consequences of differences in physique, none has attracted greater attention than the phenomenon of criminal behavior. The theory that physique and crime are closely related is usually associated with the name of Cesare Lombroso (1911), whose observations in the medical service of the Italian army led him to the belief that criminals were characterized by certain degenerative physical features, referred to as "stigmata of degeneration." These included woolly hair, a receding forehead, prominent cheekbones, a scant beard, prognathism, asymmetries of the face or body, and so on. This theory aroused considerable controversy and was attacked in particular by Goring (1913), an English criminologist, who claimed to find as many stigmata of degeneration among undergraduates at Oxford and Cambridge as in the criminal population. The only difference was that the criminals were inferior in size and weight, and this inferiority was ascribed by Goring to their lower economic position.

Most social scientists in recent years have assumed that Goring's work effectively disproved the theories of Lombroso. The belief in a relationship between physique and crime has been revived, however, in an extensive investigation by Hooton (1939), who reports the results of physical examination of almost 18,000 inmates of penal institutions in comparison with about 2,000 noncriminals. In general, the criminal was found to be smaller in size, inferior in weight and body build, with a smaller head, straighter hair, narrower jaws, and so on. Hooton concludes that crime ordinarily arises from a deteriorated organism and that the primary cause of crime is biological inferiority. This conclusion has been severely criticized, mainly because the criminals in the investigation came in large measure from lower economic levels than did the noncriminal controls, and their biological inferiority may be explained at least in part as a result of their inferior opportunities for adequate physical development. Hooton's thesis may not, therefore, be regarded as demonstrated. One might place in contrast to it the statement made by Healy (1938) that, in a series of studies made by means of composite photographs of individuals well known to him and his associates, it was impossible to distinguish the delinquent and the neurotic subjects from those who showed normal adjustment.

INDIRECT RELATIONS

In the studies so far reviewed, the attempt was made, in the main unsuccessfully, to demonstrate a direct relationship between physique and mentality. It seems much more probable that in most cases the relationship is an indirect one, and that the effect of physical characteristics frequently depends upon their evaluation by the individual himself or by those with whom he comes into contact. It is suggested by Sutherland (1939) that certain physical defects and deformities occur with greater frequency in the criminal than in the noncriminal population, not because of the relationship posited by Lombroso and Hooton, but because these defects make social adjustment more difficult and increase the probability of antisocial behavior. To put it in terms of a different theoretical framework, these defects may constitute frustrations which lead to direct and indirect aggression (Dollard et al., 1939).

Such indirect effects of physical characteristics may be observed in a large variety of situations. They constitute the basis of a substantial part of the psychological system of Alfred Adler (1917), for whom the reaction to physical inferiority represents a mechanism of fundamental importance. It is not the physical deformity or ugliness of the individual that is directly responsible for the personality, but rather the attitude of the individual toward these signs of inferiority. In another culture and under different conditions these same physical characteristics might have entirely different consequences. Excessive fat, for example, is a mark of ugliness among women in our own society, and may lead to many of the reactions described by Adler. In parts of Africa, on the other hand, it is a mark of great beauty, and a girl might labor for years to attain that corpulence which would make her sexually attractive (Van Gennep, 1909).

It is clear that similar physical characteristics may result in entirely different psychological patterns under contrasting conditions. The same may hold true to some degree in the case of constitutional types. Cabot (1938) found that those of his subjects who approximated the athletic type of Kretschmer usually possessed ascendant, influential, and extroverted personalities. This was in all probability due to the fact that theirs was the preferred physique, and their personality characteristics resulted from the self-assurance which they thereby acquired. Constitution thus may be important for its indirect rather than for its direct relationship to personality.

PSYCHOSOMATIC MEDICINE

This term has been applied to the interrelationships between conditions in the organism and the dynamics of the total behavior of the individual, especially as these enter into the practice of medicine. It is known that various types of emotional response may be accompanied or followed by definite physiological changes, sometimes leading to actual diseases in the cardiovascular system or the alimentary tract, as well as elsewhere in the organism (Alexander, 1936; Saul, 1944). It is also highly probable that various organic diseases, such as tuberculosis, angina pectoris, and so on, markedly influence the temperamental patterns of the subject. The work with the electroencephalogram (Berger, 1929; Davis, 1938) bears testimony to the complexity and intricacy of psychosomatic relationships.

INDIVIDUAL APPLICATIONS

Some relationship undoubtedly exists between physique and personality, but its precise nature is not clearly understood, and too little is known with sufficient certainty to permit of extensive practical application. The theories of Kretschmer and Sheldon may have far-reaching implications, not only for the diagnosis and prognosis of mental disturbance but also for vocational guidance and other aspects of social adjustment. It would be premature, however, to apply their techniques in the present state of our knowledge. Hooton's results might have great practical significance for the handling of delinquency and crime, since social and economic approaches would be subordinated in importance to biological factors, but the methods used in his investigation fail to carry conviction. The emphasis of Adler on the effects of organ inferiority rests upon significant clinical and biographical evidence, although here too it must be remembered that physical inferiorities do not have the same influence in all cases. Because of the wide popular use which is made of physical features as a guide to personality, the psychologist emphasizes the relative scarcity of proved relationships in this field and the need for the greatest caution in the practical application of all of these methods.

At the same time, there are a few practical suggestions which a psychologist may legitimately offer. The available material on psychosomatic medicine indicates that not infrequently temperamental aberrations, depressions, and anxieties may be the cause of organic disease as well as its product. Every individual should therefore realize the importance of adequate care and attention, not only for his physical but also for his mental well-being. In a society which places a premium on physical attractiveness the wise individual, male or female, will not neglect physical appearance and will see to it that all unnecessary defects are removed. The person of inferior stature may comfort himself to some extent with the knowledge that there is no relationship between size and ability. Even a hunchback may be helped by knowing that intellectual qualities are not usually harnessed to bodily defects.

As the individual looks at himself and at his friends he will do well to remember that the judgment of personality from physical characteristics is unwise and unfair; and that almost all varieties of physique are compatible with the highest mental and moral achievements.

READING REFERENCES

Adler, A., *The Neurotic Constitution*. New York: Dodd, Mead & Company, 1917.

Allport, G. W., *Personality*. New York: Henry Holt and Company, Inc., 1937.

Hooton, E. A., *Crime and the Man*. Cambridge, Mass.: Harvard University Press, 1939.

Hoskins, R. G., *The Tides of Life; The Endocrine Glands in Bodily Adjustment*. New York: W. W. Norton & Company, Inc., 1933.

Hunt, J. McV. (ed.), *Personality and the Behavior Disorders*. New York: The Ronald Press Company, 1944, Chap. 8, "Physiological Effects of Emotional Tension"; Chap. 17, "Constitutional Factors in Personality"; Chap. 33, "Electroencephalography."

Klineberg, Otto, *Race Differences*. New York: Harper & Brothers, 1935.

———, *Social Psychology*. New York: Henry Holt and Company, Inc., 1940, Chap. 15, "The Problem of Personality"; Chap. 20, "Social Factors in Delinquency and Crime."

Kretschmer, E., *Physique and Character* (Eng. trans.). New York: Harcourt, Brace and Company, Inc., 1925.

Paterson, D. G., *Physique and Intellect*. New York: Century House, 1930.

Sheldon, W. H., and S. S. Stevens, *The Varieties of Temperament; A Psychology of Constitutional Differences*. New York: Harper & Brothers, 1942.

———; and W. B. Tucker, *The Varieties of Human Physique*. New York: Harper & Brothers, 1940.

19. GUIDED LEARNING

T. R. McCONNELL, PH.D.

Dean of the College of Science, Literature, and the Arts and Professor of Educational Psychology, University of Minnesota. Formerly: Professor of Education and Psychology, and Dean, Cornell College; Member, President's Commission on Higher Education; President, American Educational Research Association.

G. LESTER ANDERSON, PH.D.

Professor of Education, University of Minnesota. Formerly: Director of Student Teaching and Principal of the University High School, University of Minnesota; Educational Consultant, US Office for Military Government in Germany; Lecturer, University of Michigan and Harvard University.

Learning has been defined as changes in behavior which are correlated with experience or training. Learning occurs when a person must react to a situation for which his previously acquired modes of response are inadequate. It is therefore essentially a process of adjustment, or a means of overcoming obstacles to the satisfaction of a need or motive (Melton, 1941). Learning involves much more than the accretion of specific associations. In its more complex forms it is a goal-directed process which necessitates the organization of responses in a sequence and a pattern appropriate to the end in view. The organizational factor in learning is one of its most important features. It is essential, for example, for a learner to perceive the significant relationships in a situation or task. Then it is necessary for him to organize his responses in accordance with this structure.

Human beings possess little behavior that is not the result of modification of previous behavior. Consequently, it would be difficult to find an activity in which learning does not take place. Much of this change occurs without formal tuition or guidance. The psychologist, however, is primarily interested in knowing how learning progresses under controlled conditions, such as those provided by the school, by industry in an apprenticeship program, or by business in its in-service training activities. Whether in school or out, whether in a bank or in the factory, certain general principles of learning operate, which constitute the basis for its guidance.

THE LEARNING PROCESS

After analyzing a large number of experiments in learning, Dashiell (1935) outlined the process as follows: (1) The subject must be motivated. (2) A field or complication of motives exists. (3) Obstruction is offered to the principal motive. (4) Hyperactivity, or hypertonicity, is aroused. (5) The response is multiple and varied. (6) The response is to relations of stimuli. (7) The important relation is between means and objective. (8) Selection or least action occurs. (9) The selected responses originally occur fortuitously. (10) The effects of responses are crucial. (11) The rate of learning varies in degree from gradual to abrupt.

This brief analysis indicates that the human organism, under appropriate stimulation, is an active creature striving to attain certain goals, either immediate or remote. During successive efforts to attain these ends, the individual makes progress by the perception or discovery of means-ends relationships. Learning in any task that is fairly complex is characterized by selection and elimination of responses in terms of their relevance, or lack of relevance, to the goal, and by the organization of the appropriate responses into an effective performance. In other words, the learning process is characterized by (1) variable and adaptive behavior of the individual as he tends toward the goal, (2) selection of appropriate responses in terms of means-end relations, and (3) integration of selected responses into smoother or more efficient and economical patterns of behavior.

The broad outline of an effective method of learning is therefore clear. One should first define the outcome or determine the nature of the goal as definitely as possible. Then by attempting to discern the essential features and relations in the situation, and by bringing relevant past experience to bear upon the task, he should approximate the desired performance as nearly as possible or formulate a hypothesis for solution as sagaciously as he can. This trial should be evaluated in terms of its outcomes; in other words, in terms of its relevance to the goal, its correspondence to the standard which has been set, or the extent to which it meets the demands of the problem. The next attempt should be adapted in the light of this appraisal so that it will more closely approximate the final result. Throughout the learning process, one should constantly strive for an effective organization of the performance, or of the ideas necessary for solution of a problem, since the patterning of behavior is essential to efficient operation.

MOTIVATION

Motives have three essential functions in the learning process (Melton, 1941): (1) they energize the organism or make it active; (2) they direct the activity of the organism; and (3) they have a selective function in that they determine which responses are selected and which are eliminated.

Although motives have their basis in the organic drives of the organism, they are subject to almost infinite elaboration and modification. Motives, therefore, not only stimulate learning, but are constantly being acquired or modified by learning. One of the most important purposes of education is to lead the individual to develop new interests on the basis of old ones and to establish more appropriate and more mature goals to guide his learning activity. One of the most difficult but essential tasks in teaching is to stimulate a desire to learn. One way of doing this in industry is to give the worker an appreciable part in planning and evaluating the activities in which he engages (Roethlisberger and Dickson, 1934). This participation has often proved more effective in stimulating or maintaining output than changes in wages or hours or physical conditions of work. Wherever learning is to be done, it pays to spend time and energy in creating an interest in improvement—a will to learn.

OUTCOMES OF LEARNING

Learning results in changes in behavior. The products, or outcomes, of learning take a variety of forms. Changes are brought about in our ways of looking at things, called attitudes and appreciations; in our ways of using our muscles, i.e., motor skills; in our ways of attacking mental tasks, i.e., problem solving and thinking; in our basic store of information, called knowledge; or in our understanding of symbols, called meanings. These categories overlap greatly. One would have great difficulty in classifying what has been learned in any given situation as any one learning product. In many instances, learning activities result in the concomitant acquision of knowledge, skill, appreciations, and attitudes (Bode, 1940).

TRANSFER OF TRAINING

One of the most practical problems in any educational enterprise is to apply what is learned under new conditions. This is transfer of training. It is seldom possible to teach an individual to make every response he may be called upon to perform after the period of systematic education. Transfer of training occurs in greater or lesser degree. Training may transfer to new situations, but it is seldom automatic and it usually is less in amount than one would wish. It has been generally held that transfer occurs from one situation to another to the extent to which the two situations are similar, or to the extent to which they have common elements. Broadly speaking this is true, but it is necessary to think of identical elements, not in terms of relatively independent parts of the whole, but in terms of similar relationships. In other words, transfer is most likely to take place, first, when the learner understands the essential meaning of what he learns, and, second, when he probes the meaning of the new situations which he meets.

GUIDANCE IN LEARNING

Learning usually depends upon practice conducted under favorable conditions. But uncritical descriptions of learning confuse practice with sheer repetition. It is this overemphasis on repetition which has made education in so many classrooms little more than a process of formal drill. A second fundamental principle of learning is that acts are modified by their consequences. This principle became widely known as the law of effect. Unfortunately, the effect of responses became identified with the pleasure-pain tradition in psychology, which was responsible for the notion that responses which were pleasant or satisfying were learned, while those which were unpleasant or annoying were eliminated (Gates, et al., 1948). The key to efficient learning from this point of view, is to do a thing over and over under conditions in which the correct responses are rewarded and the incorrect responses are punished. Also, an effective teacher is supposedly one who can gather together an abundance of drill material, for example, games, contests, prizes, progress charts, and other forms of extrinsic incentives, and who can keep the pupils practicing the desired responses by making the activity pleasant. There can be little doubt that some teachers have become highly skillful in conducting drill. In many classrooms, in fact, tolerable repetition becomes the maid-of-all-work in learning.

ADAPTIVE LEARNING

When practice or drill is equivalent to the repetition of specific responses to specific situations, it may well be considered a form of training, or the acquisition of specific habits. If these relatively stereotyped responses remain attached to the specific details of the situations in which they were originally learned, and the individual does not develop an attitude that they are applicable in other situations, little transfer may be expected. On the other hand, when the goal of systematic learning is adaptability, when the student generalizes his learning so that it is no longer tied to one set of situations but is available for use in a succession of appropriate situations, and when the individual consistently tries to apply his learning to new tasks, his experience may become truly educative. When practice is interpreted as a series of progressive approximations to a standard or goal of performance rather than the sheer repetition of responses, learning becomes essentially adaptive. From this point of view, the consequences which are effective in learning arise from the functional relationships of means and ends (Bode, 1940). Responses are selected and learned, not because they are pleasant in the sense of inducing a favorable feeling tone, but because they are the means of attaining the learner's goal or of satisfying his motives. Of course, it is ordinarily satisfying or pleasant to make progress toward one's objective. But it is not the feeling tone itself which is the crucial factor. The important condition is that the individual should perceive the instrumental character of his reactions.

EFFICIENT LEARNING

Learning is most efficient, therefore, when (1) persistent stimulation in the form of a task or problem evokes active efforts to adjust; (2) the goal or end-result is clear-cut; (3) the individual evaluates each attempt to reach the goal by measuring his performance against a standard or a criterion of success; and (4) the learner adapts his next trial in the light of his appraisal of the previous one.

Since efforts to learn vary greatly in their effectiveness, both teacher and learner should be aware of economical methods. Among the factors which are important and which may be controlled to a considerable extent are the following: (1) motivation, (2) knowledge of the correct performance, (3) evaluation of specific trials, (4) knowledge of progress, (5) length and distribution of practice, (6) amount and difficulty of material, (7) organization of material, and (8) organization of effort. Unguided efforts to learn to typewrite, to play golf, to acquire a foreign tongue, or to attain skill at bridge may cause so much trouble that they send one scurrying for tuition or guidance. The value of guidance in learning has been clearly demonstrated (Davis, 1945).

ACQUISITION OF SKILLS

A first essential in learning an act of skill is to know the character of the effective performance. This ordinarily cannot be ascertained by a cursory examination of the act. In fact, research frequently is necessary to define the exact nature of a complex skill. For example, the actual motor responses involved in reading were only revealed by careful investigations of eye movements during the reading process. A person who would learn an act of skill must become thoroughly familiar with the essential characteristics of the performance.

DIRECT OBSERVATION

The observation of performances and of models is a common and useful aid to learning. By watching an act performed, one can often get a clearer notion of the goal than could be secured from a verbal description. By observing an adequate performance, a standard for evaluating attempts to reproduce it may be established. However, casual observation of the performance or model is seldom effective in attaining a clear knowledge of what is demanded. One must attend to the essential cues and relationships as well as to specific details. Learning by observation is difficult.

MANUAL GUIDANCE

The effectiveness of manual guidance, that is, putting the learner mechanically through the performance, has been investigated experimentally, but the results are by no means conclusive. Gates and Taylor (1926) studied the tracing of letters on tissue paper placed over letter forms that were clearly visible through the paper. Using an experimental and control group of children without previous experience in writing, the investigators found direct practice in actual writing was more effective than tracing in learning letter forms. However, an investigation conducted to determine the relative efficacy of free and manually guided maze learning, when equated in terms of time, favored guided learning, particularly when retention was measured (Waters and Ellis, 1931).

If manual guidance gives the learner insight into the structure of the performance as a whole, or enables him to get the feel of the real activity, it is probably beneficial. Many forms of mechanical guidance, however, do not provide a realistic trial of the skillful act in a normal setting and are therefore of little value. This probably accounts for the inefficacy of tracing in the Gates-Taylor experiment.

WHOLE-PART METHOD

There is a notion that to master an act of skill as a whole one must first master the parts, and that it is relatively easy to put the parts together into a well organized performance after they have been learned piecemeal. Although it is undoubtedly true that one can first learn the elements and then combine them, the organization of the parts is often the most difficult stage. Music is more than scales, and sewing is more than a combination of particular stitches. Since the pattern of an act of skill is its most significant characteristic, it is often more economical in the end to practice the performance as a whole from the beginning.

It may be beneficial in the later stages of learning, or in striving for precison, to isolate elements from the whole and practice them singly. Similarly, when a particular part of the performance gives unusual difficulty, practice on that specific phase may be sound as a preventive or remedial measure.

WORK SITUATION

The secret of economical learning in motor skills, as in other abilities, seems to be that a thing should be learned as nearly as possible in the form and in the setting in which the performance must eventually be done. One group of students learning to typewrite were taught by a method in which, from the beginning, the practice material was comparable to that which typists use in practical work. A second group began practice on isolated and meaningless combinations of letters, and then progressed to words, phrases, sentences, and finally to business letters and other composition forms. At the end of the experiment, the first group was significantly superior in average number of letters written per practice period (Barton, 1926).

VARIABILITY OF TASK

The most significant thing that can be said about an act of skill is that it is not a series of discrete movements run off mechanically and invariably, but a flexible, integrated behavior pattern which will change, within limits, with the perceived demands of the situation at hand. Compensation in performance is nearly always demanded by the varying conditions under which an act of skill is performed. Consequently it would be unwise to practice an act under invariable conditions which require exact duplication of the performance. Much as the seeing-eye dogs are taught first to obey and then to disobey, the essential or basic form of an act of skill is learned and the details of the performance are modified to meet the varying conditions under which it must be done. The golfer probably never swings twice in exactly the same way, but adjusts his stance and stroke to conditions of the turf, the direction of the wind, the lie of the ball, and the distance to the green. A reader modifies the act of reading in terms of his purpose, the nature of the material, and the time at his disposal.

IDEATIONAL LEARNING

Organization has been said to be the most important characteristic of learning. It is much easier to learn material which has some kind of form or meaningful organization than to learn something in which the items have no intrinsic relationships. Experience with ideational or verbal material in particular bears out this fact. For example, nonsense syllables are more difficult to memorize than sentences; consecutive discourse is learned more easily than disconnected sentences; a series of numbers haphazardly arranged is more difficult to learn and recall than a list of numbers arranged according to a certain scheme or principle which the individual understands. In every case the material which is easier to learn is more highly structured. Therefore, whenever one has something to learn, he should immediately search for principles of form or organization in the material. Units of work should always be constructed around an organizing principle, and the students should search for the intrinsic relations, the structural pattern, or the meaningful sequence of ideas in managing each learning enterprise.

WHOLE AND PART LEARNING

The efficacy of learning by wholes or by parts is one phase of the broader problem of the role of form in learning. The term "whole" should be interpreted in a qualitative rather than a quantitative sense. There is little advantage in studying a body of material as a unit unless it possesses a substantial degree of internal organization. It is this factor of meaningful unity which constitutes a qualitative whole. If the material is well structured, whole methods of learning are likely to be more efficient than part methods. This statement will prove true, of course, only if the learner grasps the relationships and keeps them clearly in mind as he works at the task. The evidence on the efficacy of whole versus part learning appears on the surface to be inconclusive (McGeoch, 1931). This may be because the subjects have been required to learn by whole methods when the material was a whole only in a quantitative sense. The procedure in other instances probably failed to produce superior results because the subjects did not apprehend and utilize the organizing scheme in the material. The most effective method of learning depends upon the degree of integration, difficulty, and length of the material, and upon the age, intelligence, background, and study habits of the learner. The general principle to follow is that one should work with the largest body of meaningful material that he can understand and manage as a unit.

PRACTICE TECHNIQUES

Two study or practice techniques have been subjected to experimental investigation to determine their relative efficacy. These are (1) reading versus recitation in learning and (2) massed versus distributed practice. In studying, one can read and reread the material for the entire learning period or he can combine reading with recitation or the attempt to recall. The evidence is fairly conclusive that recitation is advantageous (Woodworth, 1938). Recall is not particularly useful, however, unless there has been enough previous reading time to develop understanding of the material (Peterson, 1944).

That distributed practice is more effective than massed practice is agreed upon almost unanimously. Perhaps in the earliest stages of learning and with very easy material the advantage is with massed practice, but with material which is more frequently encountered in normal learning activities, spaced practice will prove more economical (Woodworth, 1938).

MEANING

The most significant problems in learning are those concerned with meaning or understanding. A high school pupil may learn in chemistry that H_2SO_4 is sulphuric acid and yet be unconcerned when some of it is spilled on his clothing. The significance of H_2SO_4 is not that it is sulphuric acid but that it is an extremely active and destructive reagent. The understanding of H_2SO_4 can be expanded to include the practical knowledge of its uses in industry, how it is manufactured, the theoretical knowledge of the structure of the H_2SO_4 molecule and what makes it behave as it does.

Meaning depends upon relationships. Among the most important relationships to emphasize are those between cause and effect, between means and end, between an act and its consequences, and between some thing or idea and its uses. One of the most significant goals a person can set for himself is to learn to discover relationships, to systematize experiences, to acquire knowledge as a means of attaining purposes, and to consider the bearing of one fact, principle, or idea upon another.

FORGETTING

Recent experimentation on forgetting indicates that passage of time or disuse may not be primarily responsible for it. Rather, forgetting is the result of an active process of interference or of interaction between what is learned at one time and what is learned subsequently (Swenson, 1941). This interference has been called retroactive inhibition. The evidence indicates that interference is at a maximum when the interpolated activities which intervene between the original learning and its recall (1) are somewhat similar to but not identical with the original material; (2) are immediately adjacent to either the original learning or its recall; and (3) are themselves not too well learned. Interference is also at a maximum when the original material is not well learned or when either the original or the interpolated material is not in itself highly integrated or organized. Retroaction takes place with either verbal or motor activities. There is an increasing amount of evidence to the effect that one remembers better what he really understands.

LIMITATIONS TO LEARNING

Individual differences in organic conditions, general and special aptitudes, interests, motives, attitudes, and previous achievement have a profound effect upon learning. This has been conclusively demonstrated. Nevertheless, the influence of these factors in learning is constantly being ignored. Restrictions upon what a learner can do are set both by constitutional or native characteristics and by previous experience. Because ways to change the constitutional factors have not been discovered, educational activities must be selected with reference to them. Goals must be adjusted to individual differences in intelligence. It is important that a person take previous experience into consideration in learning. The first stage in learning, therefore, may be to make an inventory of aptitudes and experiences. Only when a person knows his capabilities, achievement, special talents, interests, and health, can he devise the activities which may be effective in changing his behavior in the desired directions.

ORGANIC STATES AND LEARNING

Certain organic states facilitate or inhibit learning. Fatigue probably inhibits learning. What passes for fatigue, however, is often lack of interest or motivation. If drives are powerful enough the individual will continue to react and to learn even after long continued activity. Studies have shown that prolonged but highly motivated mental labor decreases efficiency little although the work may become very dissatisfying. Mild emotional states probably facilitate learning, while powerful emotions have a disruptive effect. Fear, worry, and similar states are unfavorable to learning (Prescott, 1938).

PHYSICAL AND SOCIAL SURROUNDINGS

Persons with learning difficulties are often exhorted to find a place for study which is free from distractions: where lighting is right, sound is reduced to a minimum, and temperature is at from 68° to 70° F. This advice is good, but such conditions are not absolutely essential for learning. Woodworth (1938) states that "with young adults, easily put on their mettle, the result [of experiments on distraction] is usually that the distractors do not distract, except perhaps for a short time while [the subjects are] becoming adjusted to the situation." However, if motivation is not strong enough or the distraction is more interesting than the task at hand, for example, swing music on the radio while the college lad is studying chemistry, there may be a disruption of the learning activity. Whether the effect of presence or absence of others on learning is favorable or unfavorable probably depends upon such conditions as distraction, emotional states, and motivation. These factors should be considered when either socialized or individualized learning is planned.

POINT OF VIEW

The general principles of learning summarized above should not be looked upon as laws of learning in the sense that they operate invariably. Learning products are the outcomes of a complex interaction between the individual and his environment. The learner's wants, goals, ongoing activities; his attitudes and mental sets; his previous experience and the inherited limits of his nature—most or all of these may be acting in a given learning process. The kind of situation or task which the individual faces, the nature and stubbornness of the obstacles which he meets in satisfying wants or attaining goals, these and many other aspects of the environment complicate his learning. These numerous conditions which are related to the course of learning, and to economy in acquiring new ways of responding, make it clear that individual differences are as influential in learning as in any other phase of human behavior. Hence, the principles of learning are not presented as laws, but to show that they express a certain predictability of learning behavior and that they indicate some of the most significant considerations in the control of the learning that people do at all ages and in manifold situations.

READING REFERENCES

Gates, A. I., et al., *Educational Psychology*. New York: The Macmillan Company, 1948.

Kingsley, Howard L., *The Nature and Conditions of Learning*. New York: Prentice-Hall, Inc., 1946.

McGeoch, John A., *The Psychology of Human Learning*. New York: Longmans Green & Co., Inc., 1942.

Melton, A. W., "Learning," in W. S. Monroe (ed.), *Encyclopedia of Educational Research*. New York: The Macmillan Company, 1941.

National Society for the Study of Education, Nelson B. Henry (ed.), *Forty-first Yearbook: Part II. The Psychology of Learning*. Bloomington, Illinois: Public School Publishing Company, 1942.

20. VOCATIONAL AND EDUCATIONAL SATISFACTION

ROBERT HOPPOCK, PH.D.

Professor of Education, Department of Guidance and Personnel Administration, New York University, and Editor, *Occupational Index;* Diplomate in Counseling and Guidance of the American Board of Examiners in Professional Psychology. Formerly: Field Secretary, National Vocational Guidance Association; Assistant Director, National Occupational Conference, and Vocational Counselor, Board of Education, Rahway, New Jersey.

DONALD E. SUPER, M. A. (OXON.), PH.D.

Associate Professor of Education, Teachers College, Columbia University; Consulting Psychologist, American Institute for Research; Diplomate in Counseling and Guidance, American Board of Examiners in Professional Psychology; Certified Clinical Psychologist (New York). Formerly: President, Division of Consulting Psychology, American Psychological Association; Major, Air Corps, Chief, Psychological Branch, AAF Regional and Convalescent Hospital, Miami, Florida, and Research Officer, Psychological Research Unit No. 1, AAF Training Command; Associate Professor of Educational Psychology and Director of the Personnel Bureau, Clark University; and Director, Cleveland Guidance Service, Cleveland, Ohio.

We all want a reasonable degree of satisfaction. Sometimes we think life would be wonderful if we could have everything we want. At other times we realize that that would mean satiety, a state of existence only temporarily pleasant. What is best for us, probably, is a condition somewhere between complete satisfaction and complete dissatisfaction, one that would leave us enough unrealized desires to keep us at work with active interest, but would relieve us of an urge to run away from our present troubles.

It is encouraging to note that evidence from many occupations indicates that most persons, perhaps two-thirds, are satisfied with the jobs they hold (Hoppock, 1935). They would welcome an increase in salary and perhaps a promotion to the job next above theirs, but they are not worrying much about it and would rather remain in their present job than change to a completely different line of work.

CAUSES OF SATISFACTION AND DISSATISFACTION

There are a number of possible causes of satisfaction that may be kept in mind if one would be satisfied or would help others to be satisfied. These causes operate together in varying combinations.

BEHAVIOR

The characteristic reaction with which we respond to unpleasant situations is one of the most important determinants of our satisfaction, whether on the job, in the classroom, on the campus, with our friends, or in the family group. Most of us have seen how differently two persons react to the same situation. Faced with unpleasant circumstances, some persons will whine and blame their families, or their friends, or the economic system, or the world at large; others will accept these situations as the inevitable challenges of life and proceed either to adjust to them or change them if they can.

FRUSTRATION. There are three types of behavior in frustration situations where one's purpose is opposed: (1) the extrapunitive, in which the subject blames his failure on his tools, on people, or on the social situation; (2) the intropunitive, in which the subject feels that his failure to accomplish a desired end is his own fault; and (3) the impunitive, in which the subject feels that neither he nor any other person is to blame and takes his failure with a fatalistic philosophy. A group of psychologists and sociologists at Yale University have developed the thesis that frustration is the cause of most aggressive reactions (Dollard et al., 1939), which means that we must look to frustrating situations as the cause of much vocational and other dissatisfaction.

HUMAN RELATIONS. One's present behavior is influenced by experiences in early childhood. A good set of parents can be a great help in the development of normal attitudes, not only because of the heredity with which they endow their child, but because of the environment which they provide. Women stenographers who had lived a happy childhood and whose relationships with their parents had been pleasant were found to be better satisfied with their work than were those whose childhood experiences were not so happy (Quayle, 1935). An analysis of case histories showed that the people who had made the best vocational adjustments had grown up in homes in which there was more family loyalty and harmony, had less antagonism toward parents, brothers, and sisters, felt less rejection, and had more adequate family lives as adults, than those who were not well adjusted in their vocations (Friend and Haggard, 1946). This finding is what one would expect in view of the demonstrated importance of parent-child relationships in emotional adjustment.

Attitudes toward instructors in college and supervisors in industry may have their origin in early childhood experiences, for attitudes toward parental authority are frequently transferred, as the child grows older, to parent substitutes, such as teachers, policemen, instructors, employers, judges, and others having authority. Evidence reported by Vetter and Green (1932) suggests that radical and rebellious attitudes on religious and social questions are frequently caused by unsatisfactory relationships with parents. A boy who develops an attitude of genuine respect for his father may, other things being equal, grow into a student who likes his instructors and a man who is loyal to his employer. Overidentification with superiors leads to emulation of the superiors, so that the attitudes of one's instructors in college or of one's superiors on the job are accepted uncritically as one accepts the attitudes of one's parents. A normal degree of identification causes one to accept conditions more or less as they exist, working for gradual improvement. Rejection of one's superiors results in being resentful and rebellious, desiring immediate reform even at considerable cost to oneself and to others.

Roethlisberger and Dickson (1939) have shown in the Hawthorne Works of the Western Electric Company that few things have more immediate effects upon job satisfactions than relations with the boss. One manager may disrupt and shatter the morale of his workers, while another keeps his staff happy and efficient under the same material conditions. Industrial relations workers are coming to realize more and more that a favorable emotional climate in which to work is fundamental to job satisfaction, and that this depends largely on supervisory relations. Lippitt (1940) has experimentally demonstrated that democratic relationships between boys and leaders in boys' clubs are more productive of satisfaction, and also of efficiency, than are autocratic and anarchical situations. Unacceptable attitudes are responsible for more discharges from business and industry than any other noneconomic factor (Brewer, 1927). These attitudes often center on and are the product of human relations, particularly in taking and in giving orders. An intensive study in a department store clearly focused attention on the role of relationships with associates at work and with one's family (Anderson, 1929). Hale (1939) has shown that unsatisfactory family relationships are frequently the underlying cause of educational and social maladjustment and Moreno (1934) and other sociometrists have demonstrated that the facilitating of association with congenial persons results in improved group and individual morale.

Some persons who have never learned to get along happily with other people find happiness only when they can work in comparative isolation. Others find their greatest happiness in the friendships which they make among their associates. A few seem to find compensation in friendships off the job for unhappy relationships at work. Those who use realistic techniques in social relations are better satisfied than those whose behavior is not so well oriented to reality (Miller, 1940).

Studies of workers' emotions have revealed that pleasant or unpleasant family relationships have much to do with happiness on the job (Hersey, 1932). Hoppock (1935) and Miller (1940) have shown that people are more likely to be satisfied with their jobs when their families regard them as successful.

RELIGIOUS ATTITUDES. Religion is a powerful force in channelizing attitudes. It may lead one to accept all ills with fatalistic resignation, to crusade militantly for drastic change, or to work tolerantly with others toward a common goal. Inability to reconcile oneself even temporarily to inevitable disappointment or disillusionment creates many problems in mental hygiene and may send some persons into schizophrenia. Unfortunately, there has been little objective study of the role of religious attitudes in human adjustment, although it has been the subject of much inspirational writing.

SUCCESS

Success is widely regarded as one of the conditions which should bring satisfaction. But success is a relative matter. Achievement representing a high degree of success for one who started at the bottom could readily represent comparative failure for one who began at the top. Children of parents who work in factories often want white-collar jobs; merely obtaining modest positions as stenographers or clerks may therefore satisfy their ambitions, even though their net earnings are not so great as they might have been in a factory.

The higher one goes in the occupational scale the greater the amount of job satisfaction (Hoppock, 1935). The percentage of satisfied men increases with higher occupational levels of a socio-economic scale (Super, 1939). This fact suggests that absolute achievement is important in gaining vocational satisfaction for various reasons: because of a widely accepted material standard of success in our culture, because of the greater challenge to ability at the higher occupational levels, and because of greater freedom of action at these levels.

Relative success is also important in gaining vocational satisfaction. Super (1939) showed that men in occupations lower than the one in which they had previously been engaged were more likely to be dissatisfied with their work than were those who had maintained their highest attained occupational level.

SOCIAL PRESTIGE

The relative status of the individual in the social and economic group with which he identifies himself affects the social prestige of a worker and thus his satisfaction (Hoppock, 1935). The mere possession of any job, even of work relief, may be a source of great satisfaction if all of one's friends are unemployed (Morgan, 1937). A steady job may carry social prestige sufficient for satisfaction if one's associates are employed in seasonal industries. But if one's social group is composed of professional workers or business-owners the required level of accomplishment rises proportionately. One may find satisfaction in a lower job only if he can regard it as a stepping-stone or as a social contribution more significant than the material achievements of others. Failing this he may gain social prestige by distinguishing himself in an amateur activity (Super, 1941). Bakke (1933) found that the English unemployed preferred the dole with the hope that they might be reemployed in their former occupations to employment at lower occupational levels, despite financial stringency and a restricted social life.

SECURITY

Employment security affects job satisfaction in proportion to the estimate which the worker himself places upon it, rather than in proportion to the material level of security. Mr. Milquetoast in the most secure of all positions might still wish for greater security. This is not to disparage employment security. Its effect upon satisfaction can be tremendous. Hall (1934) found among engineers that the safer a man felt in his job the higher the level of his morale, which may be interpreted as job-satisfaction.

FEAR

Fear is capable of destroying confidence, dissipating ability, and demoralizing a working force. Hence it is a cause of considerable dissatisfaction. It is frequently without objective foundations, especially when strong enough to have disorganizing effects, but is associated with emotional insecurity and therefore with lack of confidence in oneself and in others.

Fears often originate in childhood or adolescent experiences. Recent and current experiences are also important in causing fear, as studies of the effects of threatened unemployment and the loss of means of livelihood have shown (Hall, 1934). They acquire additional importance because they often reactivate the residual fears of earlier experiences. Emotional maladjustment is frequently found in dissatisfied workers (Fisher and Hanna, 1931).

Persons living at home, in families which are intact, have better morale than others when unemployed, probably as a result of continued experiences of emotional and material support (Rundquist and Sletto, 1936).

Fear in combat has been studied by military psychologists, with results which have implications for civilian situations. Shaffer (1946) obtained anonymous responses to questions concerning their backgrounds and combat reactions from more than 4,500 combat flyers. Most flyers experienced fear on combat missions. The delayed reactions to these unusually strong stimuli included fatigue, restlessness, depression, instability, and loss of appetite and interest. Idleness and inability to do anything about danger increased fear. Uncertainties and unknown factors in a situation also aggravate fear. On the other hand, group spirit, doing something about the situation, the example of others, and personal pride help control and reduce fear. In business, industry, and labor unions these are factors partly within the control of leaders. From the point of view of the individual, it is well to remember that it is not so much being afraid or not being afraid that matters, as how well one is able to control fear.

MORALE

As fear can weaken, so morale can strengthen the satisfaction of the worker in his work. There are irritations and disadvantages in all occupations and institutions. Unless the worker can feel that they are worth enduring, he is not likely to be satisfied with his job. Wages, salaries, and profits are offered as one compensation. Confidence in management provides another, which in many instances is even more effective in deriving work satisfactions (Roethlisberger and Dickson, 1939; Shaffer, 1946). This is probably because confidence in management increases feelings of belongingness and emotional security.

ABILITY AND INTEREST

Abilities and interests in relation to work are causes of job satisfaction and have received more attention than any other causes.

College personnel officers are familiar with many instances of failing students who have done satisfactory and sometimes brilliant work after being transferred to courses for which they were better adapted by intelligence or special aptitude. Especially in elementary schools, experience has shown that problem students are often guilty of nothing more than a desperate protest against being expected to do work that is either above or beneath their capacities. Much of the guidance and personnel work in school and college is directed toward the correction of disparities between the abilities of the student and the requirements of the educational program. Evidence such as that gathered in the Speyer School experiments (Gates, 1936) is convincing in its proof of the importance of school work suited to the abilities of the student in keeping him, not only satisfied, but well adjusted emotionally.

Industrial personnel work seeks to achieve reasonable harmony between the abilities of the worker and the requirements of the job by careful selection of employees, assignment to appropriate activities, and such upgrading or reassignment as may be desirable. That this is important to job satisfaction has been shown definitely by the Hawthorne study, where job placement was mentioned in roughly two-thirds of 6,800 interviews. Employees placed in jobs appropriate to their intelligence levels are likely to be more satisfied than others (Anderson, 1929).

INTERESTS. Discrepancies between the interest of the man and the nature of the work are a primary cause of dissatisfaction with one's job. The better the placement, in terms of suitability of work with respect to interests, the greater the resultant satisfaction and the better the emotional adjustment (Super, 1939). Job dissatisfaction and emotional disturbance are likely to be found together in situations which involve strain and difficulty (Kornhauser and Sharp, 1932). Super (1941) has shown that men who believe they chose their jobs because they were interested in them derive more satisfaction from their work than from their avocations, whereas the converse is true of others. Men whose vocations resemble their major avocations are better satisfied with their work, change occupations less often, and are less likely to want to change occupations. Strong (1943) reported that men with interests like those of successful life insurance salesmen were more likely to sell enough insurance to make a living than men whose interests were not like those of life insurance salesmen. This suggests that congenial work is likely to result in greater efficiency. That this was not found true of teachers and aviators is probably due to the predominant importance, in these occupations, of factors other than congeniality. A study using Strong's Vocational Interest Blank found that adults who complained of occupational dissatisfaction tended to have interest patterns which were not congruent with their occupations (Sarbin and Anderson, 1942). It seems safe to conclude that other things being equal, the provision of a vocational outlet for a dominant interest is likely to result in vocational satisfaction.

Vocational guidance stands in a sense between educational and industrial personnel work, helping individuals to make the transition from school to employment and to find their way into occupations in which they can be both reasonably efficient and reasonably well satisfied. Vocational guidance, scientific placement, good labor-management relations, anything that increases the harmony between the work and the worker, helps to increase the satisfaction of the individual in his work.

SATISFACTION CYCLES

There is a rather widespread belief that young people are more likely to be dissatisfied with work than are older people, and that the latter are more satisfied with life in general. We are all familiar with the stereotype of the fiery young radical who settles down after he gets older, gets a job, has a family, and gains some perspective on life's problems. If this idea is valid, we would expect to find a larger percentage of dissatisfaction with work and social institutions among young people than among older persons.

AGE. Several investigators have considered the relation of age and satisfaction and agree that job satisfaction increases with age (Hoppock, 1935; Super, 1939; Miller, 1940). Age, however, appears to be a correlate of other causes of job satisfaction and not a cause in itself (Super, 1939). As men get older many rise in the occupational world, and the resulting freedom and feeling of success cause them to become more satisfied with their jobs. Child (1941) suggested that the reason college graduates now in the forties are better satisfied with their jobs than younger men and women may be because they graduated in time to get established before the depression, which implies that job success and not age is the cause of the satisfaction.

The development of job satisfaction may be cyclical. Men in the age group 20 to 24 tend to be satisfied, those in the age group 25 to 34 are dissatisfied, and older men are satisfied, with a temporary decrease in satisfaction at ages 45 to 54 (Super, 1939). This fits in with Buehler's theory of life stages, the suggestion being that the younger men feel satisfied to have made a start, that they then become dissatisfied not to have progressed more rapidly, and later, having either become established or resigned, they are satisfied again.

CONSTANCY. Satisfaction in one's work and social life varies from time to time and fairly consistently, but with irregularities resulting from current experiences. Each individual has his own constant cycle, ranging in length from 3 to 9 weeks, with lesser ups and downs within the week and even within a given day (Hersey, 1932). Particularly important in creating, accentuating, or prolonging cheerful moods are social participation and successful accomplishment of the day's work; in creating depressed moods, illness and anticipation of work to be done are important. Each individual has his average satisfaction level around which he varies relatively little in a given job situation.

TECHNIQUES OF MEASUREMENT

Vocational and educational satisfaction can be evaluated by a variety of techniques. These are useful to psychologists, educators, and personnel workers, and to students and employees who have occasion to consider their own vocational adjustments. Their use generally requires the aid of a psychologist or vocational counselor.

QUESTIONNAIRES AND RATING SCALES

The satisfaction questionnaire is the most common method of evaluating vocational satisfaction in large-scale investigations. It asks the respondent to answer yes or no to several questions concerning his liking for his job. These answers have been used to classify subjects as primarily satisfied, dissatisfied, or indifferent in studies by Hersey (1933), Morgan (1937), and Super (1939, 1940). Other studies (Hall, 1934; Hoppock, 1935; Rundquist and Sletto, 1936) have weighted the items in the questionnaire by a statistical technique to yield a quantitative score of degrees of job satisfaction. Some of these questionnaires incorporate rating scales for self-ratings of job satisfaction, which are also used independently.

CLINICAL INTERVIEWS

Clinical interviewing is a valuable method of studying job satisfaction (Hoppock, 1935; Roethlisberger and Dickson, 1939). According to the latter a nondirective or free association technique applied to the work situation is most effective in diagnosing and stimulating job satisfaction as it provides catharsis and results in a feeling of belongingness and even in insight on the part of the subject.

OBSERVATION

Direct observation of job satisfaction was used by Bakke (1933) who lived among the English unemployed and by Lazarsfeld and Zeisl (1938) in Austria in making an investigation of the relative satisfaction of the unemployed. The latter placed a participant observer in an unemployed village to keep records of the actions, statements, and activities of its inhabitants. In the extreme conditions which existed in this community prolonged unemployment resulted in complete apathy.

SOCIOMETRY

Moreno (1934) originated the sociometric technique in the study of personality in girls living in the cottages of an institution. The girls wrote essays about their attitudes toward their cottage mates. These were analyzed to reveal likes and dislikes. This technique lends itself especially well to the study of satisfaction in small social groups. Regrouping of social units according to congeniality increases vocational and personal satisfaction.

INCREASING SATISFACTIONS

The individual who wishes to achieve the most satisfaction in life is concerned primarily with finding specific satisfaction sources and overcoming fears that obstruct satisfaction.

FINDING SOURCES OF SATISFACTION

The first step in finding sources of satisfaction is self-understanding, followed by an understanding of one's immediate environment. Clinical experience has indicated that this is generally the necessary sequence, as lack of self-understanding may otherwise warp one's point of view. The development of an understanding of one's abilities, interests, and personality is too broad a subject to be dealt with here, but a few of the commonly used methods can be enumerated.

1. Experiences in school, avocations, social life, and on the job can be reviewed and observed, to determine which things are done most effectively and with most satisfaction. Persistent interests and enthusiasms especially deserve consideration.
2. Friends and associates are frequently a useful source of information concerning strengths and weaknesses, but it should be remembered that they are not trained observers or judges of human behavior.
3. Psychological tests vary in their value, and are often given undue emphasis in counseling. But, properly used by a trained counselor, they often provide valuable information and serve as a springboard for developing self-insight.
4. The help of a trained counselor is important because of his understanding of individual differences and of the world of work, and because the counseling relationship makes it possible for the counselee to explore his attitudes in a more detached manner than he otherwise could.

Understanding how one's abilities and interests can be used in the world of work involves going beyond self-exploration as outlined above; investigating occupations is also necessary. A few of the more useful exploratory techniques are listed below:

1. Current articles and books on occupations are listed in the *Occupational Index* and in the *Guidance Index,* published monthly by the New York University and by Science Research Associates, 228 South Wabash Avenue, Chicago, Illinois, respectively. Recent occupational pamphlets are annotated in Forrester (1946), and publishers of such pamphlets are listed. The Occupational Information and Guidance Service of the United States Office of Education, the United States Employment Service, the Bureau of Labor Statistics of the United States Department of Labor, and their counterparts on the state level, are other excellent sources of material on occupations and employment trends. The professional journal *Occupations* comments on new resources of these types as they are made available.
2. Tryouts of various degrees of verisimilitude are possible in school and college courses, clubs and hobbies, part-time and summer jobs, and volunteer work. Testing a tentative choice in such ways often helps to clarify issues and point up planning.
3. It should be remembered that the achievement of satisfaction is an ongoing process, not an isolated event.
4. Occupational information obtained by the counselee in reading, talking with people in the field, and trying out related activities should be discussed with the counselor, to insure adequate coverage of topics, comprehension of materials, and especially clarification of attitudes toward the findings so that appropriate action may in due course be initiated by the person being counseled.
5. The following list of factors affecting job satisfaction, taken from Hoppock (1935), may be useful in indicating some of the topics to be investigated

in the vocational adjustment problems of employed adults.

Check List of Vocational Factors in Job Satisfaction

The amount earned
The hours worked
Opportunities to advance
Opportunities to attain ambitions
Opportunities to help others
Opportunities to study
Opportunities to express oneself
Opportunities to experiment
The amount of independence on the job
The amount of variety in the work
The management's policies
The appreciation shown for the work
The difficulty of the work
The things done on the job
The kind of work in general
The people with whom one works
The place in which the work is done
The community in which it is necessary or possible to live
The way in which people treat the worker

OVERCOMING FEARS

Fears are difficult for the individual to deal with by himself. There are two basic ways in which fears may be overcome: (1) by removing the cause and (2) by preparing to meet whatever one fears. The former is the more difficult and time-consuming, but the more effective when properly used.

A few fundamental methods of removing the cause are listed below:

1. A clinical psychologist can help review early experiences, bring the underlying cause of fear to light, and thus objectify attitudes toward the previously emotionalized object. The reliving of these experiences in discussion, the calm examination of a difficult situation, the accepting attitude of the clinician, often reveal that the facts are not as frightening as was believed.
2. When the cause of fear is more objective, such as possible unemployment resulting from a decreasing volume of plant operations, the removal of the cause may result from taking some definite step, such as adding to one's usefulness to the firm through application and versatility, or from increasing the prospects of securing employment elsewhere, as through study in evening courses in a locally expanding field.

Since one can never remove all potential causes of fear, the more satisfied person is likely to be the one who has prepared himself to meet the difficulties which concern him most by calm analysis of their nature and by taking measures which will minimize their effect. Experience with the unemployed indicates that vocational security lies less in the job which one has than in the jobs which, as a result of one's ability, personality and reputation, one can obtain.

The following maxims may be helpful to the person adjusting to the first job:

1. Develop the habit of taking life's irritations calmly. Remember what frustrated reactions are and the tendency for moods to shift with satisfying and unsatisfying experiences.
2. Practice getting along amiably with people, whether you like them or not, especially teachers, employers, and colleagues. Remember the tendency to project your own problems, fears, and wishes onto others, and to displace early childhood attitudes from people who first caused them onto other people who seem to resemble them in some way.
3. The things you do not know or that you half-understand are more likely to bother you than the things you see clearly. Make a practice of putting your worries and fears into words, either to yourself or to an understanding person who is a good listener, and, once you really see what is bothering you, make a point of finding out all you can about the situation.
4. Remember that many problems are largely problems of attitudes, and can be cleared up by understanding, accepting, and modifying attitudes. The place to begin is with your own.

READING REFERENCES

Fisher, V. E., and J. V. Hanna, *The Dissatisfied Worker*. New York: The Macmillan Company, 1931.

Hoppock, R., *Job Satisfaction*. New York: Harper & Brothers, 1935.

Roethlisberger, F. J., and W. J. Dickson, *Management and the Worker*. Cambridge, Mass.: Harvard University Press, 1939.

CHAPTER IV

TECHNIQUES OF PERSONNEL PSYCHOLOGY

21. Job Analysis	Carroll L. Shartle
22. Personnel Interviewing	Bruce V. Moore
23. Rating	Donald G. Paterson
24. Worker Analysis	Beatrice J. Dvorak
25. Job Classification and Evaluation	Carroll L. Shartle
26. Vocational Test Construction	Herbert A. Toops
27. Criteria and Validity	Robert J. Wherry
28. Item Analysis	Robert J. Wherry
29. Personality Tests and Prediction	Edward A. Rundquist
30. Effectiveness of Selection Devices	Marion W. Richardson

21. JOB ANALYSIS

CARROLL L. SHARTLE, PH.D.

Professor of Psychology; Executive Secretary, Personnel Research Board, Ohio State University. Formerly: Chief, The Division of Occupational Analysis, War Manpower Commission; Assistant Psychologist, Member Educational Staff, Milwaukee Electric Railway and Light Company.

Job analysis is a method of obtaining information about a job. The method includes the observation of the job by a trained analyst. The particular items to be included depend upon the use which is to be made of the occupational information. A psychologist has many uses for the data obtained by job analysis. They are necessary in vocational selection, in vocational counseling, and in the development of testing techniques. Likewise, a psychologist may be concerned with problems of training, job evaluation, worker transfer, and safety. The basic information needed to aid him in the solution of such problems is best obtained by job analysis.

DEFINITION OF JOB

A job may be defined as a group of similar positions; a position being a series of tasks performed by one worker. A person may hold the position of personnel technician in an industrial firm. If there is more than one position of this type they are lumped together and called a job. There are as many positions as there are persons on the pay roll. Thus in one concern there may be three thousand positions and three thousand persons on the pay roll, but there are only three hundred jobs. In this case the average number of persons in a given job is ten.

If one analyzes and writes up each position separately the process is called position analysis. If several similar positions are analyzed together it is called job analysis. Job analysis is the more frequently used process in industry. Jobs vary greatly in the extent of the tasks performed. The jobs in small establishments usually involve a number of tasks, whereas in large productive establishments there is usually a finer breakdown and therefore a greater number of jobs for the same industrial process.

PURPOSE

Before initiating job analysis, the investigator needs to determine the specific uses to be made of the occupational information obtained. If the use of the job information is merely to identify the jobs and to designate appropriate job titles and definitions, only a brief analysis is made. If, on the other hand, the psychologist intends to develop and standardize aptitude tests and determine hiring specifications, a large number of items must be included in the job analysis.

APPLYING JOB ANALYSIS

The purpose of job analysis is to solve problems for which occupational information is an aid. All too frequently psychologists and personnel workers carry out procedures requiring the use of job information without having on hand the accurate and thorough information necessary.

DEVISING AND STANDARDIZING TESTS

Job analysis is a necessary part of the standardization of aptitude tests and trade tests. If it appears advisable to develop and standardize an aptitude test battery, those items in the job analysis which are directly related to setting up a criterion of success should be emphasized. Output standards are important here. Unless there is a reliable standard by which test scores can be validated it is not worth while to attempt to standardize an aptitude test battery. If output records are not reliable or valid indices, new standards must be developed or supervisory ratings used. It is also important in developing an aptitude test battery to consider the number of workers who are performing the same job. In some instances several similar jobs may be combined and an aptitude test battery developed for such a group of jobs. In choosing the trial set of aptitude tests, the work performed, the complexity of the job elements, and aptitudes estimated as most essential should be scrutinized. Experience indicates that the psychologist's estimate of the aptitudes required may not coincide with the actual test results obtained when the standardization is undertaken. For example, in one study analyses were made of several closely related apprentice jobs such as auto mechanic, machinist, and car repairman. A certain type of mechanical aptitude seemed important as a possible selection device. In the validation process, however, a mathematical reasoning test and certain other paper and pencil tests not usually considered to measure mechanical aptitude were much more diagnostic.

After a trial battery of tests has been validated, the results should be checked against the job-analysis data. This check will give clues to the accuracy of the judgments made in preparing the analysis.

Trade tests should be based on job analysis. Special emphasis should be given to the work performed, the complexity of job elements, and the types of machines and tools used. These job-analysis data will yield suggestions for the development of the oral, written, and picture test items. Technical terms and the sequence of operation are good sources for trade-test items. Time required to train unskilled workers should also be considered, for there is only a limited number of trade-test items in any job and in simpler jobs there are few. Therefore, if the job analysis reveals a total training time of less than six months, it is doubtful that sufficient trade-test items can be developed to make a diagnostic test. If performance trade tests are developed, the most complex job elements should be adapted to the test situation when practicable.

VOCATIONAL SELECTION

The job specification reflects those parts of the job analysis which give the information necessary for a personnel worker to have in order to select workers for employment. Frequently the job specification is placed on a card for ready reference by the employment interviewer. It is not necessary to have a complete statement of the work performed; a job summary or a definition is usually sufficient to identify the highlights of a job. Care must be taken in adapting the job-analysis information to the job specification in order that hiring requirements which really reflect the actual performance are indicated. Also, it is needless to include hiring requirements which cannot be reliably determined by the interviewer. Some job specifications include personality requirements which are not defined and which, if defined, probably could not be estimated with instruments now available.

Working conditions involved in the job should be emphasized in the analysis to determine whether women can be employed. Physical requirements should be emphasized to discover whether persons with handicaps might be employed. The job elements should be observed in the light of the physical demands that are necessary. Working conditions involving the exposure of the workers to fumes, cold, dust, moisture, and other conditions should be reviewed in terms of their probable effect on workers with various disabilities. A slight change in the job elements or working conditions will often make it possible to employ workers who have certain disabilities.

VOCATIONAL COUNSELING

Vocational counseling in a school or other private or public agency requires that the counselor have pertinent information concerning the occupational complexion of the community. From job-analysis data it is possible to prepare several types of information.

The occupational monograph is usually prepared to cover a single occupation or a group of similar occupations. A number of sources of information is used, and the monograph may run several thousand words in length, including historical information, importance of the occupation, duties, qualifications, earnings, hours, employee organizations, and places of employment.

A second type of monograph is the industry monograph showing the type of activity in the industry and at least thumbnail sketches of the principal occupations involved. If the vocational counselor is concerned only with inexperienced persons his monograph may include only those occupations in which experience is not required. It is important, however, that the more complex occupations be referred to so that one has a picture of the job sequence of upgrading in the industry. The outline for a good Occupational Monograph has been indicated (Occupations, 1940). In preparing the monograph the job-analysis data may be very brief, highlighting such items as the job title, classification title, number employed, rates of pay, working conditions, and hiring requirements. One or two typical plants in the community may have analyzed their jobs and be able to supply such information without the counselor observing the jobs himself.

The occupational description is briefer than the occupational monograph. It is based upon analyses made in several establishments.

The job-analysis data are combined and the description prepared which covers the several titles by which the occupation may be known: the classification title; work performed; tools, machines and material used; working conditions; relation to other jobs; and hiring requirements. Since the description is based upon data obtained from more than one establishment, the description does not accurately reflect the job in a particular plant but it will give the basic information necessary for the counselor. Composite descriptions have been prepared by the Division of Occupational Analysis and Industrial Services of the United States Employment Service and are available upon request for a number of industries.

TRAINING PROGRAMS

The job-analysis approach is often used in setting up training programs within the establishment and also training in outside schools. The work-performed section of the job-analysis report should be carefully studied to indicate those skills which can best be learned on the job, and to make as a part of the hiring requirements those aspects of the job which are likely to be learned by applicants in schools outside the establishment. Certain arithmetical skills, for example, might be shown in the job elements to be essential for performance. These can be learned outside the plant. On the other hand, the particular use of these arithmetical computations in terms of the job would probably have to be learned within the establishment. The relative complexity of job elements must also be considered so that training on the job can proceed from the less complex to the more complex tasks.

TRANSFER AND PROMOTION

If the problems in the establishment are concerned with transfer and promotion, the work performed as obtained in the job analysis should be compared among jobs, and the jobs grouped according to their similarities. The complexity of the job elements should be scrutinized since more complex job elements are the ones that need primary consideration in determining transfer. Similarly, if the establishment is expanding it may be necessary to break down the jobs so that for a given job the less complex elements can be performed by a worker who is less skilled. This finer breakdown of jobs, however, cannot always be made on a logical job basis because there may be insufficient work in the breakdown job to keep one worker busy continuously. Thus the place of the job in the plant process must be considered, and also the number of workers employed in the job.

IMPROVING OUTPUT

The job analysis reveals information required in the improvement of output. The sequence of job elements should be studied to see if it is efficient. The description of the work performed may be supplemented by motion pictures of job performance in order to discover where efficiency can be improved. Items in the job analysis which may reflect physical fatigue on the part of the worker can be spotted by studying the work performed, the physical demands, and the output standards.

JOB EVALUATION

The work performed, as obtained by job analysis, may be weighed, along with such factors as working conditions, hazards, and certain hiring requirements, for the purpose of establishing the rate of pay for the job. Some job evaluation schemes become overcomplex and include items which cannot be judged accurately, even by experts in job analysis. Various systems have been developed for job evaluation, some of which are discussed in Section 25.

SAFETY

Job analysis is one approach to the problem of safety. Working conditions need emphasis in determining hazards, and in studying the work performed it is possible to discover operations and sequences of operations which are hazardous and which can be improved. The machines and tools used are also important since, if a variety of machines are used in the job, accident records may indicate upon which machine the greater number of accidents occur. Also, some workers are less able to make adjustments in changing from one machine to another.

MERIT RATING

Job analysis can serve an important role in determining and describing the items to be included in a rating scale to evaluate job performance. In fact, the specific tasks in the job can be taken from the job analysis report and reworded into rating-scale items. A common fault in merit rating is that the items are too general or otherwise do not relate specifically to the tasks the worker performs.

MANAGEMENT AIDS

Job analysis can be used as an aid in management. The job information obtained from the study may reveal that certain workers are duplicating each other's work and therefore their jobs should be changed to avoid such duplication. The job analysis may show that certain essential work is not being performed at all, and such gaps likewise can be corrected. Items in the analysis which are concerned with supervision give clues to cases where the supervision is too lax or the worker oversupervised. The psychologist has considerable to contribute to organizational problems because of his knowledge of individual differences. Organization charts all too frequently are drawn up without reference to the abilities of the employees who will fill the jobs shown on the chart. For example, it may be discovered by job analysis that an executive position involves the direction of production and of public relations. If the executive chosen to fill the post is primarily trained and skilled in production, it is necessary to consider this fact in setting up the jobs of his assistants, for he probably will delegate almost entirely his public relations duties to one of his aids. One executive may request that eight or ten persons report directly to him and may have the ability to supervise each one adequately. Another executive with the same responsibilities may be more proficient if he has only one or two assistants reporting directly to him.

PERFORMING JOB ANALYSIS

THE JOB-ANALYSIS SCHEDULE

After a careful study of the use to be made of job-analysis data, a list of pertinent items should be formulated and arranged on a schedule or work sheet (Shartle, 1946). The job-analysis schedule usually begins with the job title, and the identification of the department. Following these items the schedule should list such items as work performed and the machines and tools used. These items are obtained by the analyst in the shop, on the selling floor, in the office, or wherever the worker performs his tasks. Items concerning hiring qualifications and worker abilities will be included last in the schedule since these should be based in part on the previous information which is written on the schedule. A sample job analysis schedule, designed primarily to obtain information for improving selection procedures, is shown on pages 139–140.

This work sheet was designed to obtain the kinds of information which would be important in studying a job primarily to improve selection procedures. It is noted that items regarding causes of turnover and present selection methods are included. There is also space for the analyst to write his recommendations. Supplementary sheets are utilized to expand the record of the analysis.

DEFINITION OF ITEMS

The items in the job-analysis schedule must be carefully defined. If terms are used loosely and without adequate definitions the data will be inadequate, and often misunderstanding will develop in the use of the job-analysis information. For example, one item always found in a schedule is the job title. It may be found that the pay-roll division of the establishment uses one title, the foreman uses another, and the worker, in turn, calls his job by still another name. This item, therefore, must be defined so that the analyst knows which title he is to use in identifying the job. He also must have a rule for the form of the title so that the schedules can be alphabetized; for example, should he write a title, "Fireman, Zinc Furnace" or "Zinc Furnace, Fireman"?

After the schedule items have been defined, the procedure should be tried out on several jobs and revised accordingly. Very seldom will the first draft of a schedule be the final form.

THE JOB ANALYST

Who should be selected for job-analysis work? A psychologist, an engineer, an educator, or one trained in newspaper reporting? One organization that has selected and trained several hundred job analysts reports that no one background seems best. In general the analyst must be able to express himself clearly, to understand industrial processes readily, and to have insight into the problems of the worker adjusting to his job.

Some psychologists feel that their contribution to vocational psychology is largely one of testing. However, if the vocational psychologist is to be well rounded, he should be a capable job analyst. Furthermore, job-analysis experience gives a concept of jobs and worker adjustment that cannot be obtained in any other way.

The qualifications of the job analyst must be related to the type of analysis to be performed. If the analysis is a phase in developing and standardizing tests a psychological background is necessary, although it has been found that job analysts not trained in psychology can do job-analysis work and

JOB ANALYSIS SCHEDULE

CARROLL L. SHARTLE, PH.D.

THE OHIO STATE UNIVERSITY

IDENTIFYING THE JOB: (A job is a group of similar positions, a position being a series of tasks performed by one worker.)

Title of Job: ______________________ ______________________
(As given by foreman or supervisor) (personnel department)

______________________ ______________________
(Dictionary title and code) (Title analyst considers appropriate)

Location of Job: ______________________
(Specify dept., div., unit or other appropriate designation.)

WORK PERFORMED: (What the worker does, how he does it, why he does it, with what he works and with whom he does it. Write it simply so anyone can understand it. Write concisely. Number each task separately. Follow each task with figure showing percentage of time devoted to task and a rating of task as to degree of skill as A = highest degree, B = medium degree, C = least skill. Begin each sentence with a verb and introduce each new task with a summary sentence as "Keeps time sheet records for incentive system: writes time workers begin tasks, their job rates and their payroll numbers, using time sheet provided by supervisors. Writes time of completing tasks, the names of task and gives filled in sheet covering 10 workers to file clerk—B, 30%." Get all the facts but do not pad.)

[ABOUT 15 LINES REQUIRED]

(Use supplementary sheets as required)

MACHINES, TOOLS, AND EQUIPMENT USED: ______________________
(Describe generally unfamiliar equipment, list familiar items.)

[ABOUT 7 LINES REQUIRED]

MOST SKILL REQUIRED: (Describe or refer to work performed in which most skill is required or in which workers are most apt to fail. Give reason for difficulty if possible.)

[ABOUT 9 LINES REQUIRED]

TRAINING ON JOB: ______________________
(Describe content, method, and length of training given on the job.)

[ABOUT 7 LINES REQUIRED]

SELECTION DATA:

Number employed ______________________ Turnover (average per ______________________)

Quits ______________________ layoffs ______________________ discharges ______________________

Chief causes of quits ______________________
(Be very specific; interview foremen and supervisors to find exact causes.)

____

Chief causes of discharges____
(Be very specific; interview supervisors and foremen to find exact reasons.)

____

Chief causes poor workmanship_____________________________________
(Interview foreman and supervisors to discover what qualities in workers are lacking that result in poor work by them.)

Present job requirements: Experience______________________________
(Type and Length)

Education____
(General level and specific courses)

Skills____
(Be specific.)

Personal and other____
(age, sex, personality characteristics, etc.)

Present selection methods: Interview______________________________
(Who conducts it; what is covered?)

[ABOUT 3 LINES REQUIRED]

Testing____
(What tests are used; what are the critical scores?)

[ABOUT 2 LINES REQUIRED]

Other procedures____

[ABOUT 4 LINES REQUIRED]

PHYSICAL AND MENTAL DEMANDS: (Enumerate the principal physical and mental demands in this job you consider pertinent in selection, training, transfer, use of handicapped workers, etc.; these may include such items as strength of arms, strength of legs, ability to stand on feet all day, dexterity of fingers, keenness of vision, keenness of hearing. Tell what worker does to explain the demand as "Strength of back in lifting 50 pound bags 4 feet in loading truck.")

[ABOUT 11 LINES REQUIRED]

WORKING CONDITION:____
(List and describe briefly the working conditions such as inside, outside, wet, etc., that are pertinent in describing this job.)

____

____

[ABOUT 13 LINES REQUIRED]

OTHER PERTINENT FACTS: (List any additional facts that seem pertinent in describing this job. These might include nature of supervision given, relation to other jobs, promotional sequence, etc.)

[ABOUT 11 LINES REQUIRED]

RECOMMENDATIONS: (List recommendations for improving selection, training, and other procedures or for improving working conditions, or other practices pertinent to this analysis.)

____

[ABOUT 15 LINES REQUIRED]

prepare oral-trade questions after being trained for a short intensive period in their construction. If the jobs to be analyzed involve highly technical processes, it may be necessary to select analysts for training who are familiar with the process; otherwise the training period will be unnecessarily long.

As a part of his training, a job analyst should be required to learn the background and purpose of the job-analysis program and then be given intensive training in the details of the items covered in the schedule. He should also be trained in maintaining good personal relations with management officials and workers.

The effective training of the analyst is on the job. His job-analysis reports should be reviewed carefully and he should be required to make the corrections himself. A new analyst tends to write up the job the way he thinks it should be rather than to describe actual conditions. If he is trained in psychology he may overemphasize the qualifications required of the worker and write too little on work performed and machines and tools used. If he is mechanically inclined he may overemphasize machine descriptions. If he is already familiar with certain jobs he may overemphasize these and prepare inadequate information about jobs that are strange to him. A capable job analyst usually cannot be trained in less than six months, whatever his background.

ANALYZING THE JOB

A thorough analysis requires that all sources of information be tapped. The members of the personnel department who are familiar with the job should be asked for such information in the schedule as they have available; second, the supervisor in the department in which the job occurs should be interviewed; third, the worker himself should be interviewed; and fourth, the analyst should observe the details of the job himself. The information obtained from the various sources will conflict. In general, the sources closest to job operations are the most reliable. Given the responsibility, the job analyst makes the decisions as to which information to consider valid and which to show in the schedule. If the analyst does not have this responsibility, he should show all information in his schedule with the sources, leaving the final decision to his supervisor.

A fairly close review of the analyst's work is required so that uniformity is continuously maintained. If the analysts work in crews of four or five, the supervisor of the group can be held responsible for the quality of the work. If there are several crews, each working in a different plant, the supervisors should confer frequently to compare their methods and standards.

In certain situations it may be advisable for the workers to assist in the analysis by filling in a part or all of the schedule themselves. This method gets the active participation of the workers in the program, and if well done it may reduce the time required for the analyses. In this method, however, the jobs should be observed by a trained analyst to check the accuracy of statements.

Occasionally, and especially when considerable information is wanted about a job, the analyst may actually learn to perform the tasks or he may work as an assistant to a worker whose position is being analyzed. In analyzing the job of lineman, for example, an analyst worked for two weeks as a groundman, a worker who assists the lineman. The hazards of climbing poles and handling high-tension wires prohibited the analyst from performing the job of lineman.

TECHNIQUE IN BRIEF

Job analysis is a method of obtaining job information. It includes observation of the job by a trained analyst. The items to be included in a job analysis depend upon the use to be made of the data obtained. Thus, before initiating job analysis the investigator should carefully study the uses to be made of the information and then include in the job-analysis schedule all pertinent items for which information can be reliably obtained. The vocational psychologist, even though he may not perform extensive job-analysis work, should be skilled in the technique. Such experience gives a background and understanding of the problems of vocational adjustment that cannot be otherwise obtained. While the psychologist may use job-analysis data primarily in the development of tests and in vocational selection and counseling, he should likewise contribute to problems of job evaluation, improvement of management, improvement of output, establishment of training programs, worker transfer, and safety. Finding facts about the jobs involved should be the first step toward the solution of these problems.

READING REFERENCES

Barnes, R. M., *Motion and Time Study* (2nd ed.). New York: John Wiley & Sons, Inc., 1940.

Burtt, H. E., *Principles of Employment Psychology* (rev. ed.). New York: Harper & Brothers, 1942, pp. 508–535.

Forrester, Gertrude, *Occupations: A Selected List of Pamphlets.* New York: H. W. Wilson Co., 1946.

Ghiselli, E. D., and C. W. Brown, *Personnel and Industrial Psychology.* New York: McGraw-Hill Book Company, Inc., 1948.

Otis, J. L., and R. H. Leukart, *Job Evaluation.* New York: Prentice-Hall, Inc., 1948.

Shartle, C. L., *Occupational Information.* New York: Prentice-Hall, Inc., 1946, pp. 1–95.

Viteles, M. S., *Industrial Psychology.* New York: W. W. Norton & Company, Inc., 1932, pp. 142–168.

22. PERSONNEL INTERVIEWING

BRUCE V. MOORE, PH.D.

Head of Department of Psychology, The Pennsylvania State College; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology. Formerly: Research Associate, Personal Research Federation, and Military Psychologist, US Army.

Interviewing, or face-to-face communication of person with person, is doubtless the most universally used technique in human relations. As a technical tool in clinical psychology an interview reveals far more than can be recorded by the client on a questionnaire, and usually is the medium in psychological therapy. In vocational psychology the interview is an essential procedure in guidance, employment, and employer-employee relations. In selecting candidates for jobs, and in dealing with employee maladjustment, it is important to know the effect of one personality on another, and the interview is a practical approach for determining this. Psychological practice must include the interview as a procedure to be scientifically studied and used.

INVESTIGATING THE INTERVIEW

VALID USES

The interview procedure has been found valuable in (1) obtaining certain types of information, (2) giving information or instructions, and (3) influencing or motivating people. Effective use of interviews depends on an understanding of both their valid functions and their limitations. Interviews make their most important contributions in helping to understand personal attitudes and relationships, for which other procedures, such as psychological tests, are less effective or not available. In this they often depend on the interviewee's or interviewer's estimates which are termed subjective data. Just because it is difficult to obtain reliable subjective data is no excuse to abandon attempts to establish the interview as a scientific instrument. Progress in the solution of many personal and social problems depends on it.

SOURCES OF UNRELIABILITY

Analysis of interviewing procedures that reveals the sources of errors also indicates the precautions that will guard against these errors. Brief reports of certain investigations will illustrate the most serious causes of unreliability of findings by interviewers.

INTERVIEWER BIAS. The effect of bias was revealed by Rice's analysis (1929) of results from twelve different interviewers who interviewed 2,000 homeless men with the purpose of discovering what they claimed to be the causes of their dependent predicament. One of these interviewers who was an ardent prohibitionist found the downfall of 62 per cent of the applicants to be due chiefly to liquor, while only 7 per cent had been seriously affected by economic conditions. A Socialist interviewer found only 22 per cent who attributed their plight primarily to liquor, but 39 per cent who had been affected chiefly by economic conditions. More significantly, the prohibitionist reported that 34 per cent of the applicants themselves gave liquor as the cause, and 42.5 per cent economic conditions, but the Socialist quoted only 11 per cent as blaming liquor and 60 per cent unfavorable conditions. Each interviewer was a trained and conscientious investigator, but his personal bias affected his interpretation and report, and even modified the statement made to him by the man he questioned in his particular way. A similar effect of bias was found in the results obtained by three experienced interviewers in a public opinion poll (Blankenship, 1940).

From an investigation of the interview as a basis for rating three character traits, i.e., reliability, sociability, and emotional stability, Harvey (1938) found that bias was produced by giving to the interviewer, before the interview, a character sketch of the candidate. He concluded, however, that there is some fundamental basis of reference to behavior and appearance which sets a limit to the operation of bias; for when the interviewer was given a character sketch that misrepresented the candidate as inferior in one of the traits, judgment was affected in 40 per cent of the subjects, but when true character sketches were given, there was agreement with these in 60 per cent.

A special form of bias is a stereotype or preconceived idea that persons of a specific personality type have a certain identifiable appearance, and conversely, persons of that appearance have certain traits. The unreliability in judgments of character traits resulting from classifiying the subject as belonging to a stereotype has been demonstrated by Rice (1926). For instance, 258 students were asked to identify the photographs of a European premier, a labor leader, a bolshevik, a United States Senator, an editor-politician, two manufacturers, and a financier. The bolshevik was identified by nine as a bolshevik, but by fifty-nine as the United States Senator. This was undoubtedly because the bolshevik had a wing collar, Van Dyke beard, and moustache which contributed to an appearance or stereotype that may be described as distinguished. The largest number of correct identifications was made in the case of the bootlegger who was shown in a heavy overcoat with up-turned collar, a cap, tortoise-shell glasses, and cigar gripped firmly between his lips, thus agreeing with the stereotype in the minds of many students at that time.

Form of Question. The early research of Muscio (1916) indicated that the form of a question has significant effect on the answer obtained. The change of just an article in a sentence, such as "Did you see *the* umbrella?" instead of "Did you see *an* umbrella?", affected the reliability of an answer. Before a public opinion poll, when Franklin D. Roosevelt was popular as President, The Psychological Corporation tried out the question, "Do you approve of the idea of having Thanksgiving a week earlier this year?" When the question was phrased in the form, "Do you approve of President Roosevelt's idea of having Thanksgiving a week earlier this year?" 5 per cent more persons, a reliable difference, were in favor of the proposal (Blankenship, 1940).

Interpretation of Question and Answer. Textile workers, who had gone on strike because a 10 per cent reduction in wages was announced to be summarily put into effect one week later, were interviewed (Bingham and Moore, 1941) and their answer to the question, "Are you in favor of arbitration?" was almost invariably, "No." When the workers were encouraged to elaborate or explain their answers, it was found that arbitration to many of them meant complete surrender to the employers. After the meaning of the term had been made clear to them, many stated that in general principle they were in favor of arbitration and would have accepted it earlier, but could not do so at that late stage of the strike after knowing the employers had refused arbitration until it was on a basis satisfactory to them.

LIMITATIONS OF VALIDITY

This same study of interviews demonstrated that answers to questions of objective fact, such as the date of the beginning of the strike or the number of men working in the mill, had little validity when compared to the factual record. These objective data had no present significance for the men and with no motivation to retain them and no emotional complex involved, there was no memory of the information. Subjective data consisting of the workers' attitudes on the issues of the strike and their opinions of two different rival unions checked exactly with the results of their secret union ballots and all other valid criteria. This was evidence that the interview is valid for securing certain subjective information which cannot easily be obtained otherwise and of relatively no value for obtaining objective facts in which the interviewee has no interest and which often can be gathered from documents or other objective sources.

IMPROVING THE INTERVIEW

The many variable factors affecting interviews make interviewing difficult to study and to improve. Symonds (1938) has listed these factors as (1) those inherent in the subject, (2) the interviewer, (3) the general situation in which the interview is conducted, and (4) the form and content of the interview; and he states that research in this field is difficult because of lack of standardized conditions and difficulty of recording the interview. From stenographic records of interviews he makes suggestions for securing rapport.

Under the direction of Rogers and Robinson at Ohio State University, a research program was based on phonographic recording of interviews, which was effective in improving counseling practice (Covner, 1942–44). These recordings have enabled Porter to prepare a check list of techniques for measuring counseling interview procedure. Various studies of interviewing techniques and reliability were carried on during the recent war under the Office of Scientific Research and Development, some of which used phonographic or wire recordings.

TRAINING INTERVIEWERS. Wonderlic (1939, 1942) has standardized the interview for employment and the method of training interviewers. The interview was used primarily to evaluate characteristics of verbal facility, personality, attitudes, and appearance. He selected interviewers who had (1) genuine interest in and curiosity about people, (2) intelligent understanding and sympathy, (3) experience and observation of life rather than sheltered existence, and (4) adaptability to social situations. In teaching interviewers he provided for (1) discussions of fundamental interview pattern, (2) demonstrations of correct procedure, (3) practice in interviewing under supervision, (4) repeated demonstrations of different interviews, and (5) memorization of the record forms or blanks to be used. Wonderlic states that the selection and training of interviewers resulted in markedly improved interview results, but only little improvement resulted from mere axioms, slogans, studying generalities, or casual observation. Vogler (1939) reported an analogous plan of training interviewers for public employment offices with the instruction based on job analysis. The importance of selecting and carefully instructing raters has been emphasized by Mandell (1940) for the interview or oral examination in rating and selecting candidates for civil service. He recommended that the traits rated should be based on job analysis, and should not be more than ten but a summary evaluation item is desirable.

From a summary of experimental findings relative to "dependability" Dudycha (1941) developed a number of questions which may be used as aids for an evaluation of that trait in the employment interview. In a review of research studies in employment interviewing Otis (1944) found that there was: (1) a trend toward the use of an interview guide, (2) a trend toward the use of a rating scale to record impressions gained in interviews, and (3) reason to believe that a job analysis facilitated interviewing.

EMPLOYMENT INTERVIEWS

The interview is an important part of the employment procedure and deserves adequate time for its proper conduct. For that reason, many employers provide for a preliminary interview or application blank on the basis of which there is a rough selection or screening to discover the applicants who are to be considered more carefully. Certain employers go further and collect as much data as possible, including the information on the application blank, letters from previous employers, and results of tests before having the real employment interview. The Procter and Gamble Company, for instance, administered a group of tests to college students and selected the more promising on the basis of these tests before the more time-consuming and expensive employment interview. Application blanks can yield much information, and have been standardized, the items of information validated and scored as a measure of an applicant's fitness for a job (Goldsmith, 1933, and Manson, 1925). In addition to the most complete information from such sources, however, the experienced personnel man realizes that the interview is indispensable, even though it has its limitations.

The good interviewer makes sure that the applicant has a fair picture of the position for which he is being considered. Then he seeks the applicant's cooperation in determining whether his qualifications make him fitted for the position. The applicant's personal responses, the animation he shows in thinking of the work itself, and his interest in the other employees reveal considerable about his personality and his probable adjustment to the job.

STEREOTYPES

In attempting to interpret the appearance and behavior of the applicant, the interviewer must guard against unwarranted generalizations. One form of invalid generalization is the subtle tendency to make inferences from personal appearance and physiognomy with regard to personal traits. Because the interviewer has had a pleasant or an unpleasant experience with an individual of a certain appearance, for instance, one with a hooked or Roman nose, he may react in the same manner when meeting a similar applicant in an interview. The interviewer may have developed, as a result of such experiences, fiction reading, or plays and movies, preconceived judgments as to the nature of individuals who fit in with stereotypes of his mind. This tendency to classify people into types and place all newcomers within one of these types and judge them as having fixed traits belonging always to these types is a serious source of unreliability in judgments from interviews.

THE HALO EFFECT

The error arising from stereotypes is closely related to the error of halo effect. The interviewer may be strongly affected by one or two characteristics which unconsciously cast a halo or influence over all other traits, causing him to overlook deficiencies in other aspects of the interviewee's personality (Thorndike, 1918).

SPECIFIC HABITS

Another source of error in judgments from interviews is the assumption that habits are general rather than specific. A common supposition is that an applicant who is neat in dress will likewise be neat in work, or that a person who talks rapidly and seems very much alive will also be a rapid worker, or that one with awkward physical posture will be clumsy in manual work. As a matter of fact, habits are not usually generalized to this extent, but are more frequently specific in character. Neatness in appearance should be evaluated as important in the applicant for the sake of appearance and impression on others, particularly if the position calls for someone who will meet other people frequently, but neatness of appearance frequently has no relation to neatness in such work as typing or clerical work.

RECORD OF INTERVIEW

The interviewer, particularly if he is interviewing many applicants, must record, either during the interview or immediately after it, the facts and impressions received from the interview. Most interviewers prefer not to make notes during the interview unless it is clear to the interviewee that he is recording such information as address, telephone number, or years of education and experience which he wants to be sure that he has accurately. A carefully prepared rating scale or some other device for organizing and recording the interviewer's judgment will not eliminate but will help to reduce the effect of unwarranted stereotypes, halo effects, and association of impressions with the wrong individual. In making these records, it is important to distinguish three types of information: observations of facts, statements by the interviewee, and impressions and inferences from facts. When this record is reviewed by the interviewer, or by someone else later, it is important to distinguish between actual fact and what someone thought about the person at an earlier time.

WHOLE INDIVIDUAL EVALUATION. Finally, the evaluation of the qualifications of an applicant for a particular position is not a simple summation of the individual's qualities. These characteristics of an applicant must be seen as an integrated personality functioning in a job situation. Although one trait should not cast a halo over others to the extent that deficiencies are overlooked, nevertheless it should be recognized that certain traits can compensate for lack of others. The interviewer's responsibility is to discover as accurately as possible from facts and criteria available in an interview whether the individual as a whole will successfully grow into a job situation. To help make the interview a more effective procedure in meeting this responsibility Fear and Gordon (1943) have developed an *Employee Evaluation Form*. This form is based on the assumption that the best indication of what an individual will do in the future is what he has done in the past. Past performance is not to be considered in terms of a single factor, such as work experience, but of the person as a whole.

On the basis of what is known about interviewing, skill in interviewing for the selection and placement of employees must be developed as an art in reference to the particular employment situation. An extensive survey of better practices in employment interviews in sixty-five companies was reported by Frances S. Drake for the American Management Association (1946). The employment interviewer interested in actual procedures and techniques will find suggestions in this report.

EMPLOYER-EMPLOYEE RELATIONS

FOLLOW-UP INTERVIEWS

The procedures of employment do not complete the process of adjusting an employee to his work relationships. An alert personnel man keeps informed on the records of the persons whom he places in the organization and has follow-up interviews, particularly with recent employees. This plan is important both for checking the validity of selection and for promoting employees who should be considered in filling vacancies at an advanced level. Another important aim for these interviews is that of keeping the management informed on the attitude of employees toward management. Leaders in business and industry at no time can assume that they know the attitude of those whom they supervise and that relations with them are satisfactory unless there have been avenues provided for free expression of the attitudes, preferences, and grievances of those whom they attempt to lead.

INTERVIEWS AND WORKER ATTITUDES

The very process of interviewing benefits both interviewer and the interviewed by releasing tensions and introducing a more wholesome attitude and healthy morale. It enables the employee whose views are solicited to get off his chest any injury he may have been nursing, a wholesome catharsis, which brings at once a new release of energy, good will, and zest for work. The manager or personnel representative who does the interviewing gets fresh insight into the varied patterns of human nature and their interrelationships in the working organization.

One of the most illuminating studies and a pioneer in extensive use of the interview was the outgrowth of an experiment of the Western Electric Company (Roethlisberger, 1939). It had its origin in an effort to determine by controlled experiments the effect of increased illumination on output. In order to be able to observe and measure the effect of different influences affecting individual variations in production, a small group of women workers were isolated from the other employees and special equipment was installed to record their rate of production. Not only was the lighting improved, but at successive periods later, rest pauses were introduced, a midforenoon lunch was provided, and a shorter working day was established. Following these changes in work conditions, the production record of the workers improved. The results were astonishing. Not only was the output increased greatly, but also the promptness and regularity of attendance and the general morale of the workers markedly improved. Still more astounding was the fact that when some of these conditions which were thought to be improvements were removed, the rate of production still continued upward. All this time conversations and informal interviews were held with the girls in the experiment and these revealed that the important change was their attitude toward management.

When these employers realized that the most potent of all incentives to good work was the right attitude of workers as a result of genuinely interested personnel supervision, the practical problem was to work out a procedure for developing this right attitude throughout the whole organization. The first step was to provide a new avenue of personal contact between workers and management. Sixty workers were chosen and trained as interviewers. Each employee interviewed was told that the management wanted to learn what he thought about conditions, what he liked and disliked about his work and circumstances, and anything else he cared to say. He was also promised anonymity and given entire freedom in choice of topics to be talked about. The interviewers themselves received many benefits from permitting the workers to unbosom themselves of irritations and resentments. The interviews not only supplied the safety valve for the workers, but also furnished management with valuable information of shop conditions, material, and personnel, which could at once be improved. Much valuable information was obtained which was organized into topics for discussion with the foremen and supervisors. From this whole experiment, certain main principles may be summarized: (1) The processes of a worker unburdening his feelings are in themselves a wholesome experience, impulses can be talked out instead of acted out. (2) Management can anticipate troubles before they become acute. (3) The chief function of the interviewer is to listen with manifest interest and never to argue. (4) This experience of interviewing is valuable training for foremen or prospective supervisors.

EMPLOYEE COUNSELING

The results from the interviews with Western Electric employees point to conclusions similar to those arrived at by other investigators and workers with problems of human behavior, attitudes, and adjustments. From much clinical experience and research Rogers (1942) developed a type of counseling interview which he calls nondirective or client-centered counseling because it provides the emotional freedom and motivation for the counselee to assume responsibility to feel and think through his own problem to a solution. This type of counseling has been used by Rogers and some of his students with workers in industry. Other psychologists and social workers have arrived at similar techniques in counseling; and during World War II, the many kinds of employee problems in rapidly growing shops and factories led to rapid growth of personnel counseling (War Department, 1943). Cantor (1945) evaluated these programs with particular attention to one industrial concern, and suggests some rules for the interviewer in employee counseling. In brief, these interviewers must not (1) argue, (2) give advice, (3) try to direct the conversation, (4) force answers, (5) take sides.

The interviewer must (1) listen rather than talk, (2) try to grasp what the employee does not want to talk about, (3) try to grasp what lies *behind* what the employee is expressing, (4) remain impartial and never make personal judgments, (5) above all, communicate to the speaker that the interviewer appreciates how he feels.

OTHER USES

The interview has been developed as a technique of clinical psychology, market research, and public opinion polls. Its uses as discussed here in vocational psychology illustrate the general procedures which can be adapted to other applications, such as in personal selling, journalism, medical practice, or in the work of a lawyer obtaining legal evidence. The guidance counselor, minister, and social worker must depend on interviews, though largely subjective and unstandardized. But the hazards of their use are minimized by a forewarning of the pitfalls discussed here in connection with vocational interviews. Successful interviewing requires both insight into the process and practiced skill in its use. It is not simple, but only beguilingly appears to be easy to use. Some of the other uses of interviews and scientific attempts to improve them are reported elsewhere in this book.

READING REFERENCES

Bingham, W. V., and B. V. Moore, *How To Interview* (3rd ed.). New York: Harper & Brothers, 1941.

Cantor, Nathaniel F., *Employee Counseling*. New York: McGraw-Hill Book Company, Inc., 1945.

Drake, Frances S., *Manual of Employment Interviewing*. New York: American Management Association, 1946.

Roethlisberger, F. J., and W. J. Dickson, *Management and the Worker*. Cambridge, Mass.: Harvard University Press, 1939.

Rogers, C. R., *Counseling and Psychotherapy; Newer Concepts in Practice*. Boston: Houghton Mifflin Company, 1942.

Stead, W. H., et al., *Occupational Counseling Techniques*. New York: American Book Company, 1940.

23. RATING

DONALD G. PATERSON, M.A.

Professor of Psychology and Member of Staff, Industrial Relations Center, University of Minnesota; Editor, *The Journal of Applied Psychology*. Formerly: Psychological Officer, US Army; Personnel Consultant, The Scott Company; President, American Association for Applied Psychology.

Rating is a term used in psychology to refer to the process of judging people and things. A rating scale, on the other hand, is a standardized device for recording judgments. It is a method whereby judgments may be secured in a uniform manner and treated quantitatively. Thus subjective opinions, estimates, judgments—call them what you will—can be dealt with in much the same way as test scores, performance records, and other psychological measurements.

Rating methods, broadly considered, are employed by the psychologist in a wide variety of situations. For example, they are used to find how consumers will judge the cubic content of variously shaped containers, where judges rate by the order of merit or by the method of paired comparisons (Guilford, 1936); or in the construction of an attitude scale, where the scale values of each statement may be determined by having a large number of statements rated with reference to the extent to which each reflects a *pro* or *con* attitude toward the issue in question; or in finding out what students think of the efficiency and effectiveness of a teacher, where each student records his opinion on a rating scale for judging teachers; or as a device in personnel administration for securing estimates of personality traits, where rating is probably the most widely used of all psychological methods.

PERSONNEL USES

Two reasons for such widespread use of rating methods in personnel administration may be advanced: (1) the necessity for passing judgment on the efficiency of others, and (2) the apparent simplicity of the rating scale method. It is the business of the vocational psychologist to meet the need for methods of evaluating efficiency and, at the same time, to set forth clearly the errors that may beset the use of rating scales and to describe how these errors may be avoided or at least minimized.

HISTORY

The first rating scale was devised by Pearson (1907) for securing estimates of intelligence. Rating scales came into general use following the spectacular development and utilization of Walter Dill Scott's Officers' Rating Scale in the United States Army during World War I (Rugg, 1921). The development of rating scales as used today and the extension of rating methods to other fields than personnel administration took place during the postwar 1920's. Not only were psychologists busy improving rating scales and applying them to a wide variety of situations, but critical research was undertaken to discover ways of overcoming the weaknesses of rating methods. As a result we now have a considerable body of knowledge in the form of principles and techniques which safeguard the use of rating scales.

It is well, at the outset, to understand the extent to which rating scales are now employed in personnel administration. Federal, state, and local government units require periodic ratings of subordinates by supervisors (Ordway and Laffan, 1936). Civil service commissions utilize rating scales for recording judgments of examiners who conduct oral interviews as part of the competitive examination system (Bingham, 1939). Rating scales are used in business and industry for many different groups of employees, e.g., salesmen, clerical workers, and factory workers (Bingham and Freyd, 1926). Rating scales for teachers are filled out by supervisors, colleagues, and even students (Cole, 1940). Rating scales are also used to provide an all-round picture of each individual student (Bradshaw, 1930). Finally, extensive use was made of rating methods by the armed forces in World War II (Newman, Bobbitt, and Cameron, 1946, and Murray and Mackinnon, 1946).

PURPOSES

In personnel administration rating has an educational value for those who make the ratings and those who are rated. A rater is forced to analyze behavior essential for success. A ratee, knowing that he is being judged periodically in essential traits, has an incentive for self-improvement in those traits in which he appears weakest. In addition, the scale items used give definite information as to what is regarded as important.

Rating scales provide a uniform and systematic method of expressing opinions to replace snap judgments of executives concerning workers. The process of developing a suitable rating scale for a particular position leads to agreement on the vocational requirements. Disagreements arising from emphasis on different traits are thus avoided.

Perhaps the most important purpose of rating is to provide a method whereby personnel administrators can check on the improvement of workers. Those whose development seems to be rapid or those whose ability seems to be unusual may be recognized promptly and suitably rewarded. Those doing unsatisfactory work likewise may be identified promptly, and up-graded or transferred to more suitable work opportunities. Those in need of vocational, educational, or personal counseling and those in need of special training also may be designated. Halsey (1944) has described a practical program for achieving this objective. The most recent treatment of employees rating scales is that of Smyth and Murphy (1946).

The main objective in rating is to individualize personnel administration so that individual differences will be properly dealt with. This is in contrast to traditional forms of management in which the individual is frequently neglected and even lost sight of. The contrast can be emphasized by mentioning antithetical practices: promotion by seniority *versus* promotion by merit; the spoils system *versus* civil service merit systems; pull or favoritism *versus* competency; wage adjustments on the basis of time service *versus* measures and judgments of competency.

RATING SCALES

MAN-TO-MAN COMPARISON SCALE

Each step from superior to inferior of the Man-to-Man Comparison scale is represented by an actual person well known to the rater. The rater does this work in forming his own master scale. The rater's master scale then constitutes points of personal comparison in rating other individuals. The method is similar to that used in product rating scales in which a master scale of samples is employed in judging the quality of handwriting, of projects made in vocational shop courses, and of other products such as themes, poems, and drawings.

The man-to-man scale was used extensively in judging officers in the United States Army (1918). An elaborate research report evaluating the effectiveness of the scale indicated that the method was too time-consuming, cumbersome, and difficult for the average officer to understand, and that the ratings were too unreliable (Rugg, 1921). The Scott Company, Personnel Consultants, made a thorough trial of this method in industry in 1919 but soon abandoned it. These disappointing results discredited the method, and the man-to-man comparison scale largely has been abandoned in personnel administration. Product-rating scales, however, continue to be used in a wide variety of situations.

GRAPHIC RATING SCALE

The Scott Company's Graphic Rating Method (Paterson, 1923) is simpler than the man-to-man method and makes much less demand on the rater's ability to think of abstract ranges of traits. The essential characteristics of the graphic rating scale can be observed in Figure 1 which reproduces one side of such a scale for factory workers. The traits or qualities are defined at the left, and to the right of each is a rating line subdivided by brief descriptive phrases to guide the rater in making his rating. The task of the supervisor is straightforward; he merely puts a check somewhere on each line. These check marks are then converted into numerical ratings by applying a stencil divided into ten steps. The numerical values for each of the traits are then summed to yield an over-all total rating. Figure 2 illustrates a graphic rating scale for reporting the efficiency of foremen and supervisors. Figure 3 gives the main body of the A.C.E. rating scale for students. It will be noted that space is provided for recording evidence to substantiate the rating. This important addition to a rating scale complies with the principle that evidence is even more important than judgment (Ordway and Laffan, 1935).

RICHARDSON-KUDER RATING SCALE

A modification of the graphic method has been developed which is rapidly becoming standard practice in personnel administration (Richardson and Kuder, 1933). The items to be rated, in comparison with those in Figures 1, 2, and 3, are increased in number, made more specific, and the descriptive phrases are reworded to cover specific work habits. Weighted scale values are assigned to each descriptive phrase by the Thurstone and Chave method of constructing attitude scales (1929). The average of the weighted scale values for all items checked by a rater is the rating score of a ratee. A sample part of the Richardson-Kuder Rating Scale for Salesmen is as follows:

Below are a number of statements that have actually been made about case goods salesmen. Read each statement, and put a plus (+) sign before it if it applies to this man. Put a minus (−) sign if it does not apply. Use a question mark (?) if you are not sure.

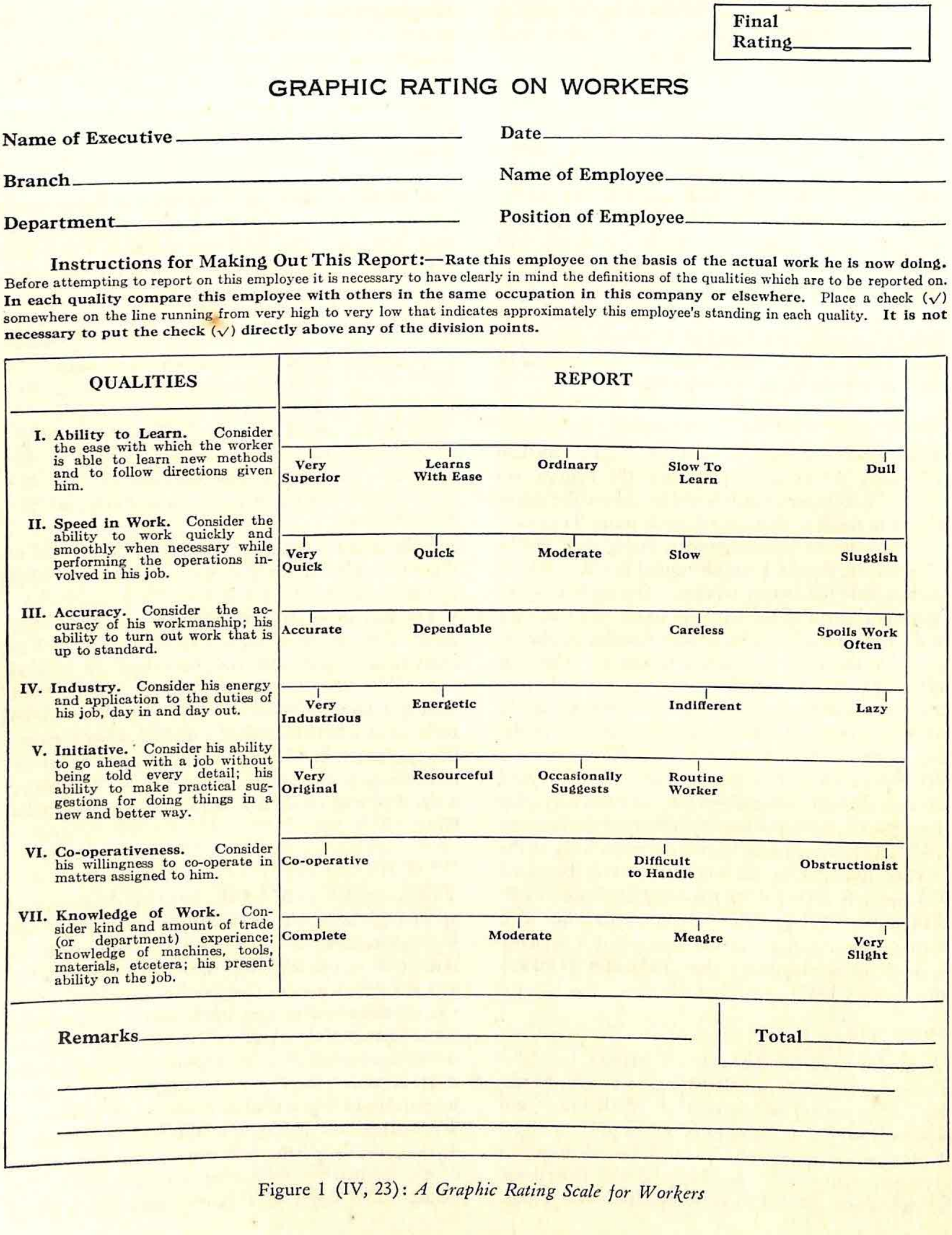
		Scale Values
. . . . .	He is somewhat in a rut on some of his brand talks	32
. . . . .	He tends to keep comfortably ahead of his work schedule . .	56
. . . . .	He is a good steady worker . .	46
. . . . .	He is weak on planning . . .	29
. . . . .	He is making exceptional progress	69

Scale values are not on the regular form used by the rater who is uninformed of the scale values of the statements.

The Richardson-Kuder technique for constructing rating scales employs the statistical method of item analysis developed in the construction of achievement tests (Hawkes, Lindquist, Mann, 1936). From among a large number of items, a smaller number of items is selected for final use. This plan increases the specificity, the discriminative capacity, and the consistency of the items comprising the final rating scale. The end result is a rating scale of high reliability.

FORCED CHOICE TECHNIQUE

This technique provides the rater with groups of apparently favorable or of apparently unfavorable items descriptive of specific behavior and forces the rater to check in each group the item most descriptive and the item least descriptive of a given ratee. Scale values, unknown to and unsuspected by the rater, are assigned to each choice. These have been determined statistically by item analysis in which the discriminative power of each item is determined by its correlation with an outside criterion of "efficiency." The method was developed by the Personnel Research Section, AGO (1946) following prior development of the method in personality testing by C. E. Jurgensen. The result is a rating procedure of high



Final Rating ______

GRAPHIC RATING ON WORKERS

Name of Executive ______ **Date** ______

Branch ______ **Name of Employee** ______

Department ______ **Position of Employee** ______

Instructions for Making Out This Report:—Rate this employee on the basis of the actual work he is now doing. Before attempting to report on this employee it is necessary to have clearly in mind the definitions of the qualities which are to be reported on. **In each quality compare this employee with others in the same occupation in this company or elsewhere.** Place a check (√) somewhere on the line running from very high to very low that indicates approximately this employee's standing in each quality. **It is not necessary to put the check (√) directly above any of the division points.**

QUALITIES	REPORT				
I. **Ability to Learn.** Consider the ease with which the worker is able to learn new methods and to follow directions given him.	Very Superior	Learns With Ease	Ordinary	Slow To Learn	Dull
II. **Speed in Work.** Consider the ability to work quickly and smoothly when necessary while performing the operations involved in his job.	Very Quick	Quick	Moderate	Slow	Sluggish
III. **Accuracy.** Consider the accuracy of his workmanship; his ability to turn out work that is up to standard.	Accurate	Dependable		Careless	Spoils Work Often
IV. **Industry.** Consider his energy and application to the duties of his job, day in and day out.	Very Industrious	Energetic		Indifferent	Lazy
V. **Initiative.** Consider his ability to go ahead with a job without being told every detail; his ability to make practical suggestions for doing things in a new and better way.	Very Original	Resourceful	Occasionally Suggests	Routine Worker	
VI. **Co-operativeness.** Consider his willingness to co-operate in matters assigned to him.	Co-operative			Difficult to Handle	Obstructionist
VII. **Knowledge of Work.** Consider kind and amount of trade (or department) experience; knowledge of machines, tools, materials, etcetera; his present ability on the job.	Complete		Moderate	Meagre	Very Slight

Remarks ______ **Total** ______

Figure 1 (IV, 23): *A Graphic Rating Scale for Workers*

GRAPHIC RATING REPORT

Name of Employee________________ Department________________

Position of Employee________________ Branch________________

Rated by________________ Date________________

Instructions for Making Out This Report:—Rate this employee on the basis of the actual work he is now doing. Before attempting to report on this employee, it is necessary to have in mind the definitions of the qualities which are to be reported on. **In each quality compare this employee with others in the same occupation in this company or elsewhere.** Place a check (√) somewhere on the line running from "very high" to "very low" that indicates approximately this employee's standing in each quality. **It is not necessary to put the check (√) directly above any of the descriptive adjectives.**

QUALITIES	REPORT				
I. Appearance and Manner. Ability to win confidence and respect through appearance, manner and general poise.	Very Attractive	Pleasing	Unimpressive	Unfavorable	Repelling
II. Industry. Perseverance, thoroughness and effective energy in daily work.	Very Energetic	Industrious		Indifferent	Lazy
III. Ambition. Desire for advancement as shown by eagerness to learn business and willingness to make the necessary sacrifices.	Very Enthusiastic and Eager to Learn		Anxious to Develop	Satisfied with Present Progress	In a Rut
IV. Initiative. Ability to get better results through originality and progressiveness in adopting improved methods.	Very Original	Resourceful	Progressive in Adopting New Methods	Occasionally Suggests Improvements	Unprogressive
V. Organizing Ability. Ability to plan work intelligently, delegate authority wisely, and make certain results are achieved.	A First-class Planner and Organizer		Plans and Organizes Work Satisfactorily	Occasional Failure in Coordinating Work	Fails to Coordinate Work
VI. Dependability. Integrity, trustworthiness, loyalty and extent to which he can be depended upon to do his best.	Absolutely Dependable		Dependable Under Ordinary Circumstances		Unreliable
VII. Judgment. Ability to compare facts and reach a sound conclusion, to estimate values correctly and to make decisions promptly.	Judgment Consistently Sound		Good Judgment in Ordinary Circumstances	Depend on Judgment of His Superior	Erratic and Unsound
VIII. Leadership. Fairness toward subordinates, ability to win their co-operation and to make them into a loyal and effective working force.	Inspires Subordinates To Best Efforts		Handles Subordinates Well	Occasional Friction in Department	Frequent Friction in Department
IX. Co-operation. Readiness and ability to join heartily with members of his department, and with other departments, for a common purpose.	Highly Co-operative	Co-operative	Sometimes Fails to Co-operate	Difficult to Handle	Obstructionist
X. Ability to Develop Men. Ability to improve his men by arousing ambition, creating interest and imparting information.	Develops Men of High Calibre		Develops Men Satisfactorily	Neglects To Develop Men	Discourages and Misinforms Men

REMARKS: (Enter the remarks needed to explain this rating)________________

____

____

Total________________

Final Rating________________

Figure 2 (IV, 23): *A Graphic Rating Scale for Supervisors*

PERSONALITY REPORT

Name of Student ______________________________

A. How are you and others affected by his appearance and manner?	Avoided by others	Tolerated by others	Liked by others	Well liked by others	Sought by others	No opportunity to observe ______

Please record here instances that support your judgment.

B. Does he need constant prodding or does he go ahead with his work without being told?	Needs much prodding in doing ordinary assignments	Needs occasional prodding	Does ordinary assignments of his own accord	Completes suggested supplementary work	Seeks and sets for himself additional tasks	______

Please record here instances that support your judgment.

C. Does he get others to do what he wishes?	Probably unable to lead his fellows	Lets others take lead	Sometimes leads in minor affairs	Sometimes leads in important affairs	Displays marked ability to lead his fellows; makes things go	______

Please record here instances that support your judgment.

D. How does he control his emotions?	Too easily moved to anger or fits of depression, etc.	Tends to be overemotional	Usually well balanced	Well balanced	Unusual balance of responsiveness and control	______
	Unresponsive, apathetic	Tends to be unresponsive				

Please record here instances that support your judgment.

E. Has he a program with definite purposes in terms of which he distributes his time and energy?	Aimless trifler	Aims just to "get by"	Has vaguely formed objectives	Directs energies effectively with fairly definite program	Engrossed in realizing well formulated objectives	______

Please record here instances that support your judgment.

Figure 3 (IV, 23): *American Council on Education Graphic Rating Scale for Students* reliability and *validity*. Furthermore, it appears to be relatively free from the ordinary errors and pitfalls in rating. The chief drawback is the extensive and expensive character of the scale construction procedure, since a special scale must be constructed for each distinct group of workers, rank and file, salesmen, army officers, and the like.

ERRORS IN RATING

Numerous errors are likely to arise in the use of rating methods. These are revealed by research studies.

ERROR OF LENIENCY

Raters judging those they know well, or for whom they feel a personal responsibility, are likely to be lenient in their rating. Raters tend to use only the above average portion of the scales in these instances. This error can be counteracted in the construction of scales by avoiding the use of extremely derogatory steps and including more steps above than below the average rating.

ERROR OF CENTRAL TENDENCY

Raters hesitate "to get out on a limb" by giving extremely high or extremely low ratings even though such ratings are deserved. They are likely to "play safe" in rating extreme deviates. This results in a piling up of ratings around the mean of the ratings for a group. The error can be minimized in construction of rating scales by increasing the number of steps in the middle ranges and eliminating extreme statements.

PERSONAL EQUATION

Some raters, like some teachers, are "hard markers," whereas others are easy. Individual differences exist among raters; a high rating from one rater may be about equal to an average rating from another rater. These differences are revealed by comparing the distribution of ratings made by a given rater with the distributions made by other raters for comparable groups of subjects. A statistical adjustment can be made which will remove this source of error (Paterson, 1923). This is accomplished by transforming ratings of each rater into a common distribution so that the ratings are made comparable (Guilford, 1936, p. 274).

HALO EFFECT

A rater usually permits his general impression to influence ratings on specific traits. The general impression swamps his judgment of details, e.g., a favorable general attitude toward a worker is likely to lead a supervisor to give him unduly high ratings in all traits even though he may be below average in several. Methods for preventing or minimizing this error are: (1) rating all workers on one trait at a time; (2) reversing the order of the descriptive phrases in chance order throughout the traits to be rated so that the rater cannot easily put down a series of checks as above or below average; and (3) training raters to be analytical by giving them practice, having all participants rate well-known individuals, and discussing with raters the reasons for divergent ratings so that they will arrive at a consensus.

SAFEGUARDS

The following principles or rules to safeguard rating procedures, which are taken from summaries of Paterson (1923), Bingham and Freyd (1926), and Guilford (1936), are designed to increase the accuracy of ratings.

1. Traits or qualities to be rated should be defined as objectively as possible, i.e., they should be described in terms of observable behavior.

2. Each trait to be rated should be restricted to a single type of activity or to the results of a single type of activity; otherwise the ratings will be ambiguous. For example, "physical qualities" if defined as physique, bearing, neatness, voice, energy, and endurance is a composite hodgepodge.

3. Descriptive phrases used to denote degrees for a given trait should not be abstract terms, such as very, extreme, and excellent. Degrees of behavior should be defined in specific terms related to the trait.

4. Traits should be described so that raters are forced to consider and report on past or present accomplishments rather than on future promise.

5. Traits should be defined in such a way as to force raters to report evidence rather than to record opinion. When opinions are called for in rating scales, additional provision for recording available evidence of the trait in question should be made.

6. Traits on which quantitative data are available should not be included in rating scales. For example, tardiness records should be used rather than ratings on promptness; actual production records should be used rather than ratings on quantity of output; and intelligence test scores should be used rather than ratings of mental alertness.

7. The method of making ratings should be simple and easily understood.

8. Ratings should be expressed in a uniform manner by all raters.

9. Ratings should be accepted only from supervisors who are competent to judge workers with a high degree of accuracy. Competence is developed in a training program conducted by the personnel worker who explains the rating procedure and gives raters an opportunity to submit preliminary ratings and to discuss them. Competence is determined by correlating each supervisor's ratings of a given group of workers with the pooled ratings of the same workers by other supervisors.

10. As many raters as practical should be utilized in securing the average rating for any given worker. A minimum of three raters is considered essential by vocational psychologists. But the more used, usually the better the average rating. Errors of different raters tend to cancel one another if many ratings of a trait are used.

11. A separate rating scale must be developed to fit each specific industrial situation. This is especially true of the Richardson-Kuder Rating Scale and the Forced Choice Technique. But even the graphic rating scale method requires the construction of different scales to cover different groups of workers, such as salesmen, foremen, office workers, and factory workers.

RELIABILITY AND VALIDITY

Reliability of ratings may be increased by training raters (Kingsbury, 1922) and by pooling the ratings of a number of supervisors. Rugg (1921) recommends a minimum of three independent ratings, whereas Bradshaw (1930) recommends from five to twenty-two judges. With properly constructed scales and with trained raters, reliability of ratings will compare favorably with other more objective psychological tests and measurements (Guilford, 1936).

Validity of ratings where an outside criterion of the same trait is available has been shown to be satisfactory (Barrett, 1914). Ordinarily, it is impossible to ascertain validity since ratings are used in those situations where an outside criterion is not available. If such a criterion were available there would be no need for ratings. Therefore, when ratings are used it is essential that every possible means should be employed to insure maximum reliability.

Bolanovich (1946) of the Radio Corporation of America has demonstrated that by the use of factor analysis and multiple correlation coefficients a fourteen-item rating chart can be reduced to a seven-item scale so that total score will correlate +0.83 with recommendations for more responsible assignments. This type of study indicates what can be accomplished by thorough statistical analysis and illustrates the fact that progressive managements are now utilizing mathematical statistics in the improvement of their personnel procedures.

If special attention is given to constructing rating scales, if the specifying conditions for securing ratings indicated here are established, and if ratings are analyzed statistically, the common errors in rating mentioned in the preceding section can be avoided and the rating scale becomes a valid and practical method for personnel administration. The method of rating will probably continue to be one of the psychological methods enjoying widespread application in all aspects of everyday life as well as in personnel administration. It is for this reason that all psychologists should be fully informed on the steps necessary to insure that ratings will be as meaningful and as accurate and as usable as possible.

READING REFERENCES

Adjutant General's Office, Personnel Research Section, War Department, "The Forced-Choice Technique and Rating Scales." *American Psychologist,* 1946, **1**, 267 (abs.)

Bingham, W. V., and Max Freyd, *Procedures in Employment Psychology.* Washington, D.C.: Shaw Publishing Company, 1926, Chap. 9, "Rating Scales."

Cole, L., *The Background for College Teaching.* New York: Rinehart & Company, Inc., 1940, Chap. 25, "The Rating of College Teachers."

Guilford, J. P., *Psychometric Methods.* New York: McGraw-Hill Book Company, Inc., 1936, Chap. 7, "The Method of Paired Comparisons"; Chap. 8, "The Method of Rank Order"; Chap. 9, "Rating Scale Methods."

Halsey, George D., *Making and Using Industrial Service Ratings.* New York: Harper & Brothers, 1944.

Paterson, D. G., "Methods of Rating Human Qualities." *Annals of the American Academy of Political and Social Science,* 1923, **110**, 81–93.

Richardson, M. W., and G. F. Kuder, "Making a Rating Scale That Measures." *Personnel Journal,* 1933, **12**, 36–40.

24. WORKER ANALYSIS

BEATRICE J. DVORAK, PH.D.

Chief, Worker Analysis Section, US Employment Service; Diplomate in Industrial Psychology, American Board of Examiners in Professional Psychology; Secretary, Division of Psychologists in Public Service, American Psychological Association; Associate Member, Panel on Personnel and Training, Committee on Human Resources, Research and Development Board of the National Military Establishment. Formerly: Research Assistant, Employment Stabilization Research Institute, University of Minnesota.

The counselor or placement worker engaged in the process of matching workers with jobs recognizes that there is a wide variation among jobs with respect to their requirements of the worker, and that there are wide individual differences among workers with respect to their qualifications for meeting those requirements. The matching process is facilitated by the use of certain techniques for estimating or measuring various worker characteristics. If the results of the use of such techniques are to be effective, the standard of reference in their development must be job requirements. Basic information for the development of worker analysis techniques is obtained by observation of the worker in his job, or job analysis.

Many human factors are important in the successful occupational adjustment of an individual and various techniques have been, or could be, devised for evaluating them. The discussion here is restricted to certain worker analysis techniques for (1) the determination of vocational interests, (2) the determination of physical capacities, (3) the measurement of occupational skills and knowledge, and (4) the measurement of aptitudes.

DETERMINATION OF INTERESTS

Considerable work has been done in the development of instruments for determining vocational interests, notably by Strong (1943), Kuder (1946), and Brainard (1945). Devices such as these yield a score which is interpreted in terms of vocational norms. Certain counseling agencies that find the scoring of such devices too time-consuming for their limited staffs use the interview for the exploration of a counselee's interests. A modification of each of these approaches is represented by the Interest Check List developed by the US Employment Service (1946). It is an interviewing aid particularly useful with persons who have no definite stated work interest and with those whose knowledge of the wide variety of tasks and activities that exist in the various occupational fields is limited. The items in this check list are based on the groupings of jobs as found in the *Dictionary of Occupational Titles,* Part IV (1944). The dictionary codes are indicated for each group of items, so that it is possible to consult the dictionary to find the lists of jobs to which a checked interest may apply. There are approximately 135 items relating to eight major fields of work, as follows: artistic work; musical work; literary work; entertainment work; clerical and sales work; service work; agriculture, marine, and forestry work; and mechanical work. The check list calls for the counselee's indication of his interests by placing check marks in front of those activities which he would like as jobs or hobbies, regardless of whether he has ever done them. There is no score to be obtained from the check list and it cannot be interpreted mechanically. The counselor reviews the check list to determine where the check marks are concentrated. This serves as a starting point from which the counselee and counselor investigate together the wide variety of occupational possibilities suggested by the various activities. The counselor discusses the checked activities with the counselee to be certain that a true vocational interest is indicated. He then consults the *Dictionary of Occupational Titles,* Part IV, to determine the kinds of jobs related to the counselee's predominant interests.

DETERMINATION OF PHYSICAL CAPACITIES

There are great differences in the physical activities and working conditions which comprise the physical demands of jobs as determined by job analysis. Certain jobs require a great output of energy in lifting and carrying heavy equipment for long periods of time; others permit the worker to sit at a bench while handling small parts. Even the same job may vary in its physical demands in different establishments because of different production methods or working conditions. There are also wide individual differences among workers in their physical capacities. Each worker may be considered to have an individual pattern of specific capacities for each of the physical activities and working conditions which are found on jobs. Hence the evaluation of a worker's physical capacities must be more specific than "pass" or "fail" on a physical examination for employment in general.

The Physical Capacities Form developed by the US Employment Servce provides for evaluation of the worker in terms of the degree to which he can perform specified physical activities and can withstand specified working conditions (1945). The items include twenty-seven physical activities, such as walking, lifting, and climbing; and twenty-seven working conditions, such as inside, outside, cold, and humid. Information on these same items is obtained in the job analysis to provide information with respect to the requirements of jobs. The physical capacities information about the worker can then be matched with the corresponding physical demands information about a wide variety of jobs to determine their suitability for his specific capacities.

The US Employment Service has also developed aids for interviewing disabled persons. These consist of pamphlets, each of which describes a disability and suggests questions which the counselor should ask in order to elicit information concerning the person's physical capacities.

MEASUREMENT OF SKILLS AND KNOWLEDGE

Various types of vocational achievement or trade tests have been developed as aids to the classification and selection of workers purported to have acquired the skills for a particular job. They include performance trade tests, oral trade tests, picture trade tests, and similar devices.

PERFORMANCE TRADE TESTS

The performance trade test is a measure of occupational skills that have already been acquired by the applicant through experience or training. The typing and stenographic tests developed by the US Employment Service, and now widely used in its local employment offices, are representative of performance trade tests for office workers.

The typing test consists of typing from plain copy for 10 minutes, and is preceded by a warming-up practice exercise of 3 minutes. The dictation exercises consist of 240 words each and are given at 80, 96, or 120 words per minute, depending upon the preference of the applicant after a trial exercise. A 20-minute time-limit is allowed for transcribing the notes on the typewriter. The norms are based on the scores of employed persons. Separate norms are provided for the interpretation of speed and accuracy. These norms meet the demands of practical business situations and have proved more useful than combined speed-and-accuracy norms.

Viteles (1932) has devised a performance trade test for motor vehicle drivers which requires the operator to drive over a standard route with an inspector who observes the operator's driving habits and records the results of observations in accordance with a standard reporting procedure. The report covers operating items, such as starting, speed, crossings, stopping, turning, backing, passing other vehicles, general economy features, changing tires, and care of equipment.

Performance trade tests probably offer a better index of an applicant's occupational skill than any other type of trade tests, but it is not always feasible to use them. Employment offices cannot install all the equipment necessary for administering performance trade tests for the large number of skilled occupations.

ORAL TRADE TESTS

Oral trade questions provide an indirect measure of occupational skill by measuring occupational knowledge acquired from work experience in that occupation. The Army trade tests developed during World War I are well known, but many of them are now obsolete. These tests were partially revised by Thompson (1936), but a thorough restandardization was not undertaken.

The US Employment Service has developed and standardized oral trade tests for more than 200 occupations. These tests are used in US Employment Service offices. They were also used by the Military Services during World War II. A typical question from an oral trade test for machinist follows:

Question: What are you doing to a stud bolt when you clamp it in the jaws of the chuck, true it up with a piece of chalk, and countersink the end of the piece?

Answer: Centering.

The standardization of oral trade questions for a job involves several steps which will be described.

FORMULATION OF QUESTIONS. A job analysis is made, and questions are devised, for the job according to the following standards: use simple wording, use the terminology of the worker, avoid catch questions, avoid questions requiring a yes or no answer, and emphasize important elements of the job with questions which elicit specific answers in one or a few words.

LOCAL VERIFICATION. The accuracy, adequacy, and difficulty of the questions are checked with workers in the occupation, and revisions are made when necessary. Then they are administered in one locality to three groups: (1) experts, (2) apprentices, and (3) workers in related occupations. The results are analyzed statistically and the best questions are selected for further tryout.

NATIONAL VERIFICATION. The selected questions are administered in various parts of the country to experts, to apprentices, and to workers in related occupations in order to obtain an indication of the general applicability of each question for the particular occupation wherever it exists. Statistical analysis of the data indicates the most diagnostic questions, which are then selected for the oral trade test for that occupation.

NORMS. Standards for the interpretation of raw scores are developed for each oral trade test used in the US Employment Service and are expressed in the following terms: (1) well informed, (2) some information, and (3) little information. The minimum score of the first category is established where there is maximum differentiation between the scores of experts and the scores of all other workers; the second category is established between this point and the score where there is a maximum differentiation between the scores of workers in related occupations and the scores of all workers in the occupation. At least 85 per cent of the experts in the standardization group make scores in the first category and approximately 90 per cent of the workers in related occupations make scores in the third (see Figure 1).

USE. The oral trade tests add to the value of the interview by providing additional data upon which the employment interviewer can base a judgment, serve as a check on the applicant's information about an occupation, and elicit supplementary information about the specific kind of experience he has had.

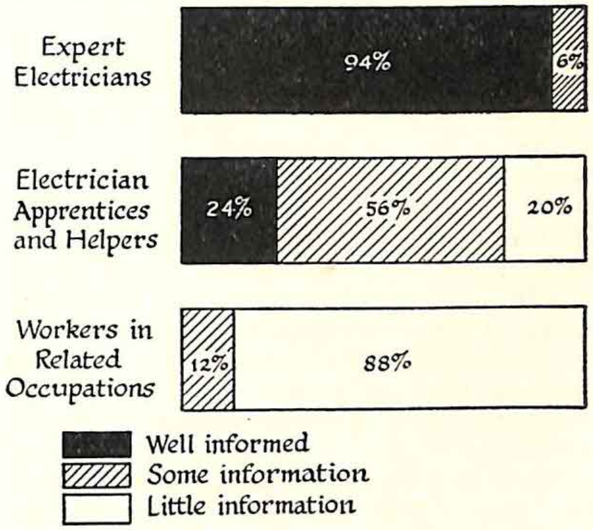


Figure 1 (IV, 24): *Differentiation of Three Groups of Workers on Oral Trade Test for Electrician, House Wiring*

OTHER TRADE TESTS

The content and standardization of written trade tests are similar to the oral trade tests just described. They can be administered to groups, whereas the oral questions must be administered individually. They are, however, less practical for use with workers in the skilled trades because they demand the ability to read and write. Furthermore, it is necessary to administer them as a test, whereas the oral trade questions can be incorporated into the interview.

Picture trade tests are similarly developed and they may be administered either in written or in oral form. Pictures of tools, machines, materials, or products are presented to the applicant who is required to identify them, or certain designated parts of them, and to indicate their uses. Blue prints may be used similarly.

MEASUREMENT OF APTITUDES

While trade tests measure a person's present occupational proficiency, aptitude test batteries measure one's potentiality for readily acquiring proficiency in an occupation.

STANDARDIZATION OF APTITUDE TEST BATTERIES FOR OCCUPATIONS

The process of standardizing a battery of aptitude tests differs somewhat from that of standardizing trade tests and involves several steps which will be described.

JOB ANALYSIS. By direct observation of the job, complete information is obtained concerning the duties of the workers employed in the occupation for which a battery of tests is to be standardized.

CRITERION. A suitable and reliable criterion, or measure of job proficiency, must be determined. The choice of a criterion depends upon the nature of the job. Link (1919) reported the use of average number of shells inspected per hour as a criterion in the development of tests for shell inspectors. For mail distributors, O'Rourke (1926) used foremen's ratings and records of speed and accuracy in distributing mail. Among the criteria used in the Minnesota Mechanical Abilities Investigation (1930) was a measure of the quality of workmanship on products made by boys in shop courses. The US Employment Service has used average hourly production in the development of aptitude test batteries for hosiery pairers and skein winders; average hourly earnings for hosiery seamers, shoe sewers, and garment sewing-machine operators; and work samples for aircraft riveters and can packers.

EXPERIMENTAL SAMPLE. A homogeneous sample is obtained by selecting persons performing the same kind of work and having approximately the same age, education, and experience.

EXPERIMENTAL BATTERY OF TESTS. The tests administered to the experimental sample are selected chiefly on the basis of job-analysis information and the results of previous studies of the same or similar occupations.

ANALYSIS OF DATA. The scores on the tests are compared with the criterion scores to determine the group of tests most highly related to the criterion (Bingham and Freyd, 1926; Guilford, 1936; Hull, 1928; Stead, Shartle, and Associates, 1940). The US Employment Service has used the Wherry-Doolittle Test Selection Method to select the combination of tests having maximum validity. Tables are established for the conversion of raw scores into weighted scores, and norms are developed for the interpretation of the sum of the weighted scores.

CHECK STUDIES. Batteries of aptitude tests developed by the US Employment Service must have general usefulness. Therefore check studies, using the same five steps described above, are made on the same occupation in plants in various parts of the country. Thus batteries are developed which can be used in the selection of persons for a given job wherever it occurs.

FOLLOW-UP STUDIES. The efficiency of a battery is determined by follow-up studies. The proficiency of test-selected workers after learning their work may be compared with their test scores made at the time of selection. The disadvantage of this procedure is that it is likely to restrict the range of both proficiency and test scores. This difficulty may be obviated by selecting persons for the job without reference to their test scores, and later determining the relationship between their test scores and their proficiency on the job; or by comparing the job proficiency of groups of test-selected and nontest-selected workers.

GENERAL BATTERIES

The US Employment Service has developed general batteries for workers employed in families of related occupations. For example, a general clerical battery was developed on samples of bookkeeping-machine operators, index file clerks, ediphone transcribing machine operators, calculating-machine operators, typists, and adding-machine operators.

TESTS FOR CLERICAL WORKERS

Considerable work has been done on aptitude testing of clerical workers. Andrew and Paterson (1934) reported that clerical workers in general were markedly differentiated from the general population by the Minnesota Vocational Test for Clerical Workers. Evidence is available indicating the usefulness of intelligence tests in predicting promotability in clerical work (Bills and Pond, 1933). The clerical battery developed by the US Employment Service consists of number-writing, arithmetic, letter-digit substitution, and number-checking tests.

TESTS FOR SALESPERSONS

Batteries of tests for the prediction of success of salespersons have been devised by Scott (1916), Goldsmith (1922), Andrews (1922), Ream (1924), Swartz (1924), Manson (1925), Craig (1925), Russell and Cope (1925), Gallup (1926), Müller (1928), Anderson (1929), Munro and Raphael (1932), Lovett and Richardson (1934), and Shultz (1935). The US Employment Service developed a battery of tests for the selection of department store salespersons after administering an experimental battery of eight tests to 260 salespersons in three department stores. Of all the tests, interest and personality items were found to be the most highly related to a composite criterion of sales production variates.

TESTS FOR MANUAL WORKERS

Pond (1927) found that the best combination of tests for selecting toolmaker apprentices was the Scovill Classification Test, the MacQuarrie Test for Mechanical Ability, and the Wiggly Block Test. Viteles (1929) developed a battery for selecting electric substation operators. Included were five paper-and-pencil tests: adaptations of the pursuit, blocks, and location subtests of the MacQuarrie Test for Mechanical Ability and adaptations of the blocks and series-completion tests of the Army Beta; and four performance tests: puzzle box, directions, learning, and persistence tests.

The US Employment Service has developed aptitude test batteries for a number of manual occupations. Some of these occupations together with the tests which are in the batteries are as follows: for transfer knitter (seamless hosiery industry), tracing, tapping, and tweezer dexterity tests; for hemstitching machine operator, arithmetic, tweezer dexterity, tracing, and finger dexterity tests; for jewelry wrapper, peg-board and finger-dexterity tests; for coil assembler (electrical equipment industry), tapping, dotting, peg-board, and finger-dexterity tests; for pairer (hosiery industry), location, tracing, number-writing, and number-checking tests; for punch press operator, paper form-board and dotting tests; for mechanical inspector, vocabulary, peg-board, and finger-dexterity tests; for inserter (printing and publishing industry), peg-board and finger-dexterity tests; and for electric range assembler, location and finger-dexterity tests.

GENERAL APTITUDE TEST BATTERY

A relatively new development in the field of aptitude testing is the General Aptitude Test Battery constructed by the US Employment Service. This battery differs from batteries of tests for specific occupations in that it is a comprehensive series of aptitude tests measuring the most important occupational abilities and predicting successful job performance in a variety of fields of work; an aptitude test battery for a specific occupation measures only the two or three most important abilities required for that occupation. The use of the specific batteries is appropriate for use in selection and placement work, where attention is focused upon the specific job opening and its requirements, and the objective is to select the best qualified individuals from all the available applicants. The counselor, on the other hand, focuses attention on the individual to help him select the kind of work which is most suitable for him and is interested in testing to explore the possibilities of various kinds of work with a single battery of tests.

The General Aptitude Test Battery consists of fifteen tests chosen as a result of a factor analysis study of a large number of tests. It measures ten occupational aptitudes: G, intelligence; V, verbal ability; N, numerical ability; S, spatial ability; P, form perception; Q, clerical perception; A, aiming; T, motor speed; F, finger dexterity; and M, manual dexterity. Of the fifteen tests, eleven are paper-and-pencil tests and four are apparatus tests. The entire battery requires about two and one-fourth hours for administration.

This battery is being standardized on representative workers employed in various occupations to obtain norms for groups of occupations established according to similarities in tested abilities. These norms are expressed as occupational aptitude patterns which consist of minimum aptitude scores required for each group of occupations. The individual aptitude profile is matched with the occupational aptitude patterns to determine the field of work that is most suitable for the person's abilities.

NEXT STEPS

During the next few years work on the standardization of the General Aptitude Test Battery will continue, so that eventually it can be scored for the entire range of occupations arranged in groups. Research will be done on the determination of additional tests to be included in the battery to measure other aptitudes. Much work remains to be done on the objective measurement of physical capacities and on personal data items and personality traits in relation to a vocational criterion.

READING REFERENCES

Dvorak, B. J., "The New USES General Aptitude Test Battery." *Journal of Applied Psychology,* 1947, **31,** 372–376.

Shartle, C. L., et al., "Ten Years of Occupational Research." *Occupations, the Vocational Guidance Magazine,* 1944, **22,** 387–446.

Stead, W. H., et al., *Occupational Counseling Techniques.* New York: American Book Company, 1940.

War Manpower Commission, Division of Occupational Analysis, "Factor Analysis of Occupational Aptitude Tests." *Educational and Psychological Measurement,* 1945, **5,** 147–155.

25. JOB CLASSIFICATION AND EVALUATION

CARROLL L. SHARTLE, PH.D.

Professor of Psychology and Executive Secretary, Personnel Research Board, Ohio State University. Formerly: Chief, The Division of Occupational Analysis, War Manpower Commission; Assistant Psychologist, Member Educational Staff, Milwaukee Electric Railway and Light Company.

JOB EVALUATION

In industry considerable attention is given to the classification of jobs into levels or grades for the purpose of arranging pay scales. Using job analysis data, the various jobs are ranked, rated, or compared, and then classified. If the position held by each worker is analyzed and classified, the procedure is called position classification. Frequently the procedure, whether it be applied to jobs, positions, or a combination of both, is called job evaluation and implies that the various positions and jobs are classified for the purpose of assigning wage rates.

PLANNING A PROGRAM

When an organization desires to initiate a job evaluation system, careful consideration should be given to planning each step and being sure that the management has a full understanding, and has given its endorsements. Many programs have bogged down because an official failed to cooperate and there was not sufficient management backing to carry the program through to completion.

COVERAGE

It is always well to consider the kinds and number of jobs which will be eventually covered in a program for this will determine in part the methodology to be employed. If there are several hundred employees performing a variety of tasks, it is better to plan one study to cover them all rather than set up several. Sometimes a plan will be developed on one group of jobs and cannot be later expanded to others because the system is too specialized. There is also the decision regarding how high up in the organization the system is to be applied. Will it cover only hourly paid employees or will it cover only jobs up to a certain annual salary?

JOB CLASSIFICATION METHODS

There are four methods which may be applied in classifying the jobs. Certain advantages are claimed for each.

GRADING AS A WHOLE

Several grades may be established on the basis of present wage levels or a modification of present levels. Each level is then defined in terms of general job performance such as "works under close supervision; does straight typing, simple arithmetical computation, and ordinary filing." The description of the job is then reviewed by an individual or a committee and is placed in the appropriate category. Careful wording of the definitions is necessary. It is best used on jobs which are all of the same general type of activity and where the number of classifications is small; otherwise the process becomes cumbersome and judgments are difficult to make. Small establishments in particular have used this method successfully.

RANKING

The jobs may be ranked from the highest to the lowest on the basis of a single criterion, such as skill required, learning time, or an over-all general criterion of job worth or job complexity. A careful statement of the criterion is necessary for the judges to follow. After the rankings are made, the jobs are regrouped into several classifications or grade levels. Wage rates are then assigned to the grade levels.

POINT SYSTEM

The most widely used method is the point system. Several factors such as learning time, responsibility, initiative, and working conditions are selected and defined. An arrangement is set up for rating how much of or to what degree each factor is required in the job. Numerical weights are assigned to the various degrees. Each job is rated according to this scale and the total points are computed. Each job then has a numerical score. The various grades are assigned ranges in scores or points, and the jobs are thus assigned to the proper grades.

This method, as will be discussed later, often utilizes factors which overlap considerably, and which have too many degrees of breakdown under each factor. Persons not understanding these limitations are likely to make extravagant claims for the system since it appears accurate because it is complex.

FACTOR COMPARISON METHOD

This method, as in the case of the point system, requires the selection and definition of several factors. Instead of assigning points, however, a group of key jobs are ranked according to the importance of each factor (Benge, Burk and Hay, 1941). The key jobs are ones which are clearly defined and whose rates of pay are not subject to controversy. The key jobs vary in salary from the lowest to near the highest of the jobs to be evaluated. The typical or average wage paid for each job is divided among the factors in proportion to the importance of that factor in the job as determined by the rankings. Each job to be evaluated is then compared with the key jobs for each factor and a value for each factor is assigned to the jobs. The total values represent the points each job receives. The total values are then fitted into a salary scale.

STATISTICAL STUDIES

It is apparent that in most job evaluation systems using factors there is considerable overlapping, and almost identical results could be obtained if the number of factors were reduced considerably. High intercorrelations among factors and between factors and total points or values have been reported in several studies. For example, two plans for job evaluation point system were studied (Rogers, 1946). In the first plan (factory jobs) there were six factors: skill, working conditions, responsibility, mentality, mental application, and physical application. Most of the variability could be accounted for by the first item which correlated 0.96 with the job grade. Skill and working conditions when combined correlated 0.97 (shrunken multiple correlation). The addition of any other variable decreased the magnitude of the correlation. In the second plan (office jobs) most of the variability could be accounted for by the first three of the following factors in the plan: initiative, training, mental application, mentality, analytical ability, personal requirements, executive responsibility, monetary responsibility and dependability. The multiple correlation between job grade and the three factors was 0.99. In each plan, one primary central factor "skill demand" accounted for most of the variance in job grade.

In a study involving a factor comparison system (Lawshe and Wilson, 1946), "skill demand" was reported responsible for 98 per cent of the final variation. The factors in the system studied were skill requirements, working conditions, mental requirements, physical requirements, and responsibility. The first three factors when weighted correlated 0.99 with the original scale.

The factor analysis studies of job evaluation systems point to two factors "skill demands" and "job characteristics." "Skill demands" includes primarily those items which the worker must possess, namely, skill, mentality, training, and initiative. "Job characteristics" are primarily in the job such as working conditions and hazards.

The statistical studies indicate that fewer factors can be used without altering the end results appreciably in terms of the grades assigned to the various jobs. It is sometimes desirable, however, to include more factors than are necessary in order to have the system accepted by executives and workers who may want certain factors included regardless of their relationship to other factors.

OCCUPATIONAL CLASSIFICATION

Groupings of closely related jobs are called occupations. There are perhaps 30,000 occupations in the United States. Over 20,000 of these have been defined and classified by the United States Employment Service in the Dictionary of Occupational Titles. The Dictionary is used as an aid in interviewing, classifying and placing job seekers who visit Employment Service offices throughout the country. It is also used by the Armed Services in noting the civilian occupations of men and women who enter or leave the services.

A second widely used occupational classification scheme is that of the United States Bureau of the Census. Occupational data obtained when the census is taken are classified and used according to this system. Although the Dictionary and the Census Classification have points in common, they necessarily vary in structure because they are applied differently.

USES IN CLASSIFYING OCCUPATIONAL DATA

The Dictionary, with its great number of titles and definitions, serves as a source for the identification of job terminology. Suppose that it is necessary to classify the respondents occupationally in making a survey by the questionnaire or interview method. Occupational titles alone are misleading; an individual may give his occupation as a clerk, while he may be a chief clerk in a large organization in a position that involves considerable executive performance or he may be a file clerk with meager responsibility. One should always ask the respondent in questionnaire or interview to state his duties and the industry in which he works, as well as the title of his job. The job title can then be looked up in Part I of the Dictionary where duties and industrial designation can be checked against the Dictionary definition so that the job may be properly identified and classified and, if necessary, assigned a code number. The code numbers are found in Part II of the Dictionary. The actual code numbers, either of 3 digit or of 6 digit, may be used thereafter in place of titles if the data are to be processed by punch-card methods. One finds sufficient detail in the Dictionary definitions so that in conducting an occupational interview for classification purposes it is possible to discuss the actual work performed with the interviewee. This makes it possible to arrive at fairly accurate classifications.

In contrast to Parts I and II, Part IV of the Dictionary contains a classification of fields of work rather than occupations. The major breakdowns are as follows:

- 0-X Professional, technical and managerial work
  - 0-X1 Artistic work
  - 0-X2 Musical work
  - 0-X3 Literary work
  - 0-X4 Entertainment work
  - 0-X6 Public Service work
  - 0-X7 Technical work
  - 0-X8 Managerial work
- 1-X Clerical and sales work
  - 1-X1 Computing work
  - 1-X2 Recording work
  - 1-X4 General clerical work
  - 1-X5 Public contact work
- 2-X Service work
  - 2-X1 Cooking
  - 2-X3 Child care
  - 2-X5 Personal service work
- 3-X Agricultural, marine, and forestry work
  - 3-X1 Farming
  - 3-X8 Marine work
  - 3-X9 Forestry work
- 4-X Mechanical work
  - 4-X2 Machine trades
  - 4-X6 Crafts
- 6-X Manual work
  - 6-X2 Observational work
  - 6-X4 Manipulative work
  - 6-X6 Elemental work

Part IV is known as the Entry Classification and is used widely for counseling and classifying persons who lack specific occupational skills or who may desire to change from one occupation to another field of work. It also contains lists of training courses and casual work experience which may have a relationship to the various fields of work.

READING REFERENCES

Benge, E. J.; S. L. H. Burk; and E. N. Hay, *Manual of Job Evaluation*. New York: Harper & Brothers, 1941.

Chesler, David J., "Reliability and Comparability of Different Job Evaluation Systems." *Journal of Applied Psychology*, 1948, **32**, 465–475.

Lawshe, C. H., Jr., and R. F. Wilson, "Studies in Job Evaluation: 5. An Analysis of the Factor Comparison System As It Functions in a Paper Mill." *Journal of Applied Psychology*, 1946, **30**, 426–434.

Moore, H., "Problems and Methods in Job Evaluation." *Journal of Consulting Psychology*, 1944, **8**, 90–99.

Otis, J. L. and R. H. Leukart, *Job Evaluation*, A Basis for Sound Wage Administration. New York: Prentice-Hall, Inc., 1948.

Palmer, Gladys L, "The Convertibility List of Occupations and the Problems of Developing It." *American Statistical Association Journal*, 1939, **34**, 693–708.

Rogers, R. C., "Analysis of Two Point-Rating Job Evaluation Plans." *Journal of Applied Psychology*, 1946, **30**, 579–585.

Shartle, C. L., *Occupational Information*. New York: Prentice-Hall, Inc., 1946.

Stead, W. H., Shartle, C. L., and Associates, *Occupational Counseling Techniques*, New York: American Book Company, 1940, pp. 175–206.

US Employment Service, *Dictionary of Occupational Titles*. Washington, D.C.: Government Printing Office, 1939.

26. VOCATIONAL TEST CONSTRUCTION

HERBERT A. TOOPS, PH.D.

Professor of Psychology and Chairman, University Intelligence Test Committee, Ohio State University; Chairman, Committee on Technical Research, Ohio College Association. Formerly: Standardizer and Assembler, Army Trade Tests; Trade Interview Specialist, US Department of Labor; Test Specialist, Camp Grant E. and R. Schools; Research Associate, Institute of Educational Research Vocational Guidance Inquiry; Consultant, Minnesota Mechanical Test Inquiry, Civilian Personnel Division of the Office of the Secretary of War, US Navy, The National Roster, and Wright Field; Inventor, Stylus Method of Self-Scored Tests; Member, National Research Council; Compiler of 19 forms of the Ohio State University Intelligence Test.

The employer, seeking to fill a position, asks, which man can do this job best? Youth, pondering a lifework, voices a complementary question, In what occupation can I succeed best? Both questions have a common element, the anticipation of maximal proficiency of a certain man for a certain job. Vocational tests may assist both.

VARIETIES OF VOCATIONAL TESTS

The most useful classification of vocational tests is on the basis of their purpose: (1) measures of achievement, progress, or attainment, primarily to aid the employer in selection of workers, and (2) prognostic tests, to aid individuals in choosing a lifework.

ACHIEVEMENT TESTS. In this category there are four principal kinds of tests classified according to the purpose for which they are used: (1) trade tests for hiring the best man for the job, (2) apprentice tests for measuring the individual's progress in learning a trade, (3) proficiency tests to determine whether a tradesman in training is ready to graduate, and (4) promotion tests to note whether an employee is ready for promotion.

PROGNOSTIC TESTS. There is a wide variety of prognostic tests each of which may be useful for only one or for a variety of vocations. Included among these tests are: (1) Mechanical aptitude tests to predict the ability of the individual to acquire a wide variety of the more complicated mechanical skills including both (*a*) performance tests, e.g., Minnesota Assembly Test, and (*b*) paper and pencil tests, e.g., O'Rourke Mechanical Test and Stenquist Mechanical Aptitude Test; (2) clerical aptitude tests; (3) interest tests, e.g., Kuder Preference Record and Strong's Vocational Interest Blank; (4) special aptitudes tests for specialized vocational fields, such as (*a*) medicine, e.g., Moss Medical Aptitude Test, (*b*) law, e.g., Yale Law Aptitude Test (*c*) engineering, (*d*) teaching, and (*e*) specialized fields, e.g., Frye's Signpainting Prognosis Test; (5) manual dexterity tests, e.g., Minnesota Manipulation Test and O'Connor Tweezer Dexterity Test; and (6) tests of mental factors of vocational significance, such as spatial visualizing, e.g., McCauley's Tetrahedron and Mann's Dynamic Cube Test.

Experts do not agree as to the functions which such tests measure. Even less may one depend on the label a test bears. Any classification at best is but arbitrary. Tests of achievement, for example, may always be employed at their proper regression weights as aptitude tests for prognosticating success in learning a trade.

TECHNIQUES OF CONSTRUCTION

The technique of building such tests consists essentially of three steps:

1. A variety of tests designed to measure trade skills or vocational aptitudes are administered to tradesmen of different degrees of skill or to a random sampling of apprentices on or before entering the occupation.

2. After some little time on the job, the occupational success of the testees is determined from independent evidence, such as work records, earnings, employer judgments, and ratings of fellow workers. When constructing aptitude tests, it is desirable, indeed, almost essential, to follow the success of the apprentices throughout their learning period, collecting evidence of their learning rates. As a result, sooner or later for each individual there is obtained an over-all quantitative success or criterion score.

3. By appropriate statistical means, ordinarily by multiple regression equations, those tests which in

combination best predict the criterion scores are selected for the final battery. The resulting test may be used to select apprentices or to give a boy seeking an occupation an adequate notion of his probable success or failure in the occupation which, it is assumed, he is considering for a life work.

Trade tests may be constructed after trade competence has been gained by the rank and file of workers employed as testees, for we are trying here to measure the results of formal and informal trade instruction. Aptitude tests ought to be administered to examinees ahead of their training in order to exclude the possibility of including in the tests elements of competence which the experienced workers gain on the job and consequently did not possess when entering the trade. To include many such elements would mean that even those of the best aptitude might be denied admission to the trade.

CRITERION SCORES

Let us assume that we are building an aptitude test. Our first and most important task is the obtaining of adequate criterion scores.

PRINCIPLES. Principles to be followed in obtaining criterion scores include the following:

1. The learners of the experimental group must have as nearly comparable work and work conditions as is practicable.
2. The measurement of progress must be continued long enough to ascertain substantially the relative ultimate success of each learner.
3. The standard of accuracy of work and the motivation of the workers must be controlled, at least to the extent of insisting upon a minimal amount of each. Particularly where the speed of machines affects the amount of work done, the rate of working must be controlled, or corrected for, with the result that the amount of work of a minimally acceptable quality done by any worker may then be taken as his criterion score.
4. The various aspects of success, particularly the important ones, the independent factors of the success profile, must be measured and included in the final composite. Judgment, following an adequate library study of previous research, and aided by the advice of practical experts is the customary means of deciding what to include in success.
5. The originally obtained scores must be reduced to comparable scores.
6. Some reasonable system of weighting the comparable scores must be employed to obtain for each individual a single success score, a unitary criterion score, which will be acceptable and accepted as measuring the relative success of the persons in the experimental group.

Certain builders of vocational tests prefer to have two or more criterion scores for each person and with different test batteries or with differential weights for the subtests of a somewhat longer battery attempt to predict each of several criteria separately, i.e., to predict, element by element, the successive scores of a criterion profile of each examinee.

The criterion is the most important single consideration in building vocational tests, albeit often slighted. Half of the funds available for a prognostic test may be profitably employed in developing it.

SCALES FOR MEASURING WORK-PRODUCTS. It may be necessary to devise objective scales to measure comparably the quality of the several work-products produced. For example, in the Minnesota Mechanical Test Inquiry (1930) cooky-cutter scales and scales of wire-splicing, made up of samples ranging in merit from very poor to very good and scaled by the same techniques commonly employed for building a handwriting scale, were nailed to a board and were used as a scale to improve the judgment of the several judges who independently judged by their aid each boy's sheet metal and electrical wiring products respectively.

SELECTING EXPERIMENTAL TESTS

Personal judgment unfortunately has to be the chief resource in deciding what tests are to be administered to an experimental population, but there are additional principles.

PREDICTION OF CRITERION. The tests should combine to advantage in the prediction of the criterion success scores. They should complement but not duplicate one another. Stated statistically: the tests to be employed should correlate highly with the criterion but low, and preferably zero, with one another. All of the mental functions or requirements of the occupation or job, as determined by a previous job analysis, should be adequately represented in the trial tests. A useful precaution is to include certain tests that measure performance versus the substitute-reaction paper-and-pencil tests. Performance tests have the merit of not reflecting to a great extent the factor of general intelligence which creeps in, willy-nilly, into paper-and-pencil tests.

Composite judgments of the merit of test-item alternatives prior to their experimental tryout is useful and economical (Kuder, 1937). More tests should be included in the tryout than are to be included in the final battery in order that the poorest may be discarded. In fact the problem is to determine the minimal number of tests, each of minimal length which, when properly weighted, will correlate maximally with the criterion. If possible each test should consist, at this state, of more items or tasks than the final edition is to contain with the expectation that, after the item analysis, the poorer items of the retained tests will also be discarded.

RELIABILITY. Each of the experimental tests should be long enough to be individually reliable. If this is not the case, the statistical constants of the test, such as validity coefficients, intercorrelations and standard deviations, will all be in substantial error. Most data on motor tests published prior to 1918 are worthless because the investigators generally did not understand the importance of reliability. Many of the earlier tests need to be lengthened from 5 to 10 or more times their original lengths.

SALES APPEAL. Finally, since vocational tests still have to be accepted in many quarters, it is well to keep in mind that any test which looks good to the prospective users will have advantages over one of otherwise equal merit which does not. A rote memory test, for example, therefore should consist of objects related to the occupation and its related fields rather than of nonsense syllables.

TRYING OUT EXPERIMENTAL TESTS

The experimental test materials have now been assembled and are ready for the next operations. All or most of the following steps usually need to be done, and in approximately the order given below.

ADMINISTRATION. For all save speed tests administer the tests by the work-limit method to a preliminary experimental group on whom a rough criterion of competence in due time will be forthcoming. The usual precautions regarding testing conditions should prevail. The test items at this stage will consist in general of free-association questions or simple trade tasks, each permitting of alternative answers or task scores. Without the recognition inherent in multiple alternatives, the aims here are: (1) to ascertain logical-to-the-examinee wrong-answer alternatives for such forms of tests if they are to be employed, and (2) to provide data for the several sorts of analyses mentioned below.

PRELIMINARY SCORING FORMULAE. Determine whether scoring formulae are necessary or desirable and, if so, ascertain the formulae required, adopting practical or rounded formulae for such as are appreciably improved thereby. Only time limit tests in general are amenable to, or are improved by, the use of scoring formulae. Economy of scoring often will require that only right answers shall be counted, particularly if electrical scoring machines are not to be employed. The use of more items, electrically or otherwise scored, often may more cheaply achieve what a scoring formula would be able to achieve with fewer items. One should not forget that the usefulness of the finished test depends materially on its simplicity of administration.

RELIABILITY. Determine the approximate reliability of each test and lengthen each as may be necessary. The Brown-Spearman reliability formula for odds-evens is appropriate. The reliability of the total battery, when completed, should be at least 0.90 and preferably 0.95 or higher, particularly if used for giving individual advice and making individual vocational decisions. Accordingly the number of trials, the number of items, or the time allowance of the test should be increased if the preliminary tests do not have individual reliability coefficients of at least 0.80 to 0.85. Higher coefficients are, of course, desirable.

ITEM VALIDITY ANALYSIS. Make an item-validity analysis (Toops, 1936) of the individual responses against the criterion, eliminating items which are so lacking in validity as to show little chance for improvement when and if altered by being cast, for example, into multiple-choice form.

RECASTING ITEMS. If the test material allows, recast the promising items into multiple choice form; choosing for the right answer that alternative which has high positive validity and also agrees with authority as correct; and for wrong answer alternatives, logical wrong answers, preferably those provided by the examinees themselves in answer to the free association questions, which have large negative alternative validity coefficients.

RESCORING. If the test form allows only of free associational, recall answers construct a scoring key in which the credits of items are proportional to their alternative validity coefficients. Partial credits for partially correct over-all performances may be ascertained. A constant, of any needed magnitude, may be added to each of the allowable scoring points of a given subtest to avoid undesirable negative credits. These have no effect on correlations. Items with many omissions may be excluded unless omissions have a desirably large negative validity coefficient. Omissions are always the sixth alternative of all five-alternative multiple choice tests. Kuder (1937) has shown that the validity of an item, that is, of the right answer, is mathematically predictable from (1) the validity coefficients of omissions and of the several wrong answer alternatives, and (2) the several standard deviations of the wrong answer alternatives and omissions. It follows that, to improve any alternative or any scoring point of an item is to improve the validity of the item of which it is a part, and so, finally, of the entire test.

By means of addends patterns-of-total-response to performance test items may be ascertained (Toops, 1942). The validity of each such pattern may be ascertained by means of the item alternative coefficient, just as if the pattern itself were an alternative. Credits for such patterns may be determined and entered upon a printed scoring sheet which identifies all allowable, or all possible, patterns with their credits, for encirclement of the credits to be added for the total scale score. For example, to omit a given task may be highly symptomatic. The validity coefficient of omissions is the measure of this significance. At this stage tests and items on the doubtful edge of desirability may well be included. They can be discarded later if they do not prove valuable.

Regression Analysis. The scores and criteria having become available, subject the tests to a multiple regression analysis, possibly still further narrowing down the number of acceptable tests. The Wherry-Doolittle method (Shartle, 1940) is an acceptable technique for this.

Revision and Readministration. Prepare the revised form of the test and administer it to a large population, five hundred or more persons if possible, on whom a very carefully derived criterion subsequently will be compiled.

Revalidation of Item Alternatives. Subject the accepted tests to an item-alternative validity analysis, employing the item-analysis of the IBM scoring machine, if possible, to shorten the work. The aim here is to insure that (1) every item finally used shall have a positive validity coefficient above a minimum of, say +.20; (2) every wrong alternative of every item of a multiple choice test shall have a sizable negative validity coefficient, say from −.10 to −.25; and (3) omissions of items shall have either negative or approximately zero validity coefficients.

Items having zero, negative, or low positive coefficients of the right answers, comprising perhaps one half to two thirds of the original items, are at once discarded. Items having positive, zero, and negative-but-very-nearly-zero coefficients for their wrong answer alternatives, but fair to good positive right answer coefficients, are carefully scanned to note whether they may be salvaged by revising the offending wrong answer alternatives. It should be observed that employment of a neutral alternative, one estimated as not very likely to be chosen by anyone however poor in ability, will reduce the number of functioning alternatives of the item by one and undo the adverse effect of a wrong answer alternative with a positive coefficient.

Final Revision. The final test may now be prepared and used. As yet the test has no norms or quotable reliability or validity coefficients and its difficulty is only roughly ascertainable. However, it will possess, in maximal degree, the all-desirable validity, and it will have at least an acceptable reliability coefficient. If enough copy has been prepared, two or more alternative forms should be constructed. This will provide retesting an examinee and may serve as a basis for judgment as to the adequacy of the test-taking motivation in individual instances. If now the test is administered by the work-limit method, which usually is the method favored by the examinees, a subsequent item analysis of the printed-test items will reveal the identity of the most promising items to be chosen from the final form or forms for inclusion in still later forms to insure that the validity of the latter also will be high, even surpassing that of the initial forms.

STANDARDIZING THE FINAL TEST

Norms. Norms will accrue if the users of the test report their distributions on standard distribution form-sheets accompanied by directions stating carefully what kinds of examinees are allowable as standardizing subjects. Publishable reliability and validity coefficients also will grow out of the various researches of the several users. To be useful, all these reports should be rendered as comparable as possible, for example, odds-even reliability coefficients should be computed in general only on work-limit tests; and validity coefficients should preferably be computed only after approximate final vocational competence of the examinee has been attained and, if possible, on groups with known ranges of ability and standard social composition. If final competence has not been achieved, then all criterion scores should be taken, as nearly as possible, after identical individual lengths of practice.

Aids to Interpretation. At this stage, to make the test maximally useful to counselors, interviewers, and others, it is useful to construct probability-of-success tables, showing for each possible total test score what is the probability of attaining each land-marked degree of success in the occupation, or occupations, of which it is prognostic. Such tables may be used by the examinee himself to determine the significance of his test score. If these tables are adequately made and are sufficiently explanatory, an interpreter will be unnecessary.

PRACTICAL CONSIDERATIONS

COST. Construction of a test by all the above steps will require, at the very least, several months for their carrying out, even in the simplest occupations. The building of a valid test by such procedures, accordingly, is expensive. The Minnesota Mechanical Ability Test (1930), subsidized by the National Research Council, required about three years to build and is reported to have cost somewhere in the neighborhood of $40,000. The techniques are now so well known, however, and with the all-but-automatic machines available for scoring the tests and for doing the statistical work, it is safe to conclude that an equally adequate study might now be carried out both more quickly and much more cheaply.

NO DUPLICATION. It is well, in building tests, to administer in the first experimental stage additional highly valid tests measuring other well-known unique mental functions than the one in which we are interested, with a view to eliminating such of our trial tests and items as measure too much what is already better measured by the latter. In the preliminary phase of construction of the Minnesota Mechanical Ability Test some half or more of the originally administered tests were excluded because they correlated too highly with general intelligence. An intelligence test was included to make it possible to eliminate those "mechanical tests," which really were only mediocre "Intelligence tests" in disguise.

In this case the Toops multiple ratio method was employed. That one of about 20 tests which correlated least with intelligence, and therefore was the worst possible "intelligence test" was accepted as a possibly good mechanical test. Thereafter the intelligence score was our criterion to be predicted. We proceeded to try to make the worst possible intelligence test out of the largest possible number, properly weighted, of "mechanical" tests. To this one-worst-possible intelligence test was added that one of the remaining nineteen "mechanical tests" which when optimally weighted for predicting intelligence gave the smallest multiple correlation with intelligence as a criterion of the two-test battery. Then to this two-test weighted battery was added that one of the eighteen which when optimally weighted for predicting intelligence raised the multiple ratio coefficient least, and so on. The increments in the multiple ratio coefficient become larger and larger with the addition of better-predicting tests to the "nonintelligence" battery. When the multiple ratio coefficient rose to 0.30 the process was stopped. By this approach, a third or more of the tests were discarded as being too heavily weighted with intelligence. One could then operate on the remainder, which passed this screening, to note which ones of these nonintelligence "mechanical tests" were valuable for predicting the mechanical criterion. The tests so chosen, when weighted optimally for predicting the mechanical criterion, did not exceed a correlation of 0.20 with intelligence. The battery thus selected was really a mechanical test largely freed of intelligence.

The theoretically desirable figure here, of course, is zero; the aim being to build such a test as will correlate zero, if possible, not only with intelligence, but also with all other unique mental factors.

MAXIMIZING HUMAN RESOURCES. By means of a vocational test possessing such properties there are denotable a maximal number of salvageable persons, such as those who were low in intellect, and therefore destined to fail in academic schools, but above average in mechanical ability, and therefore quite capable of being successfully taught trades by patient teaching. The proportion of all the testees who fall in the category mentioned is at a maximum when the correlation of mechanical ability and intelligence is zero. This salvaging of skilled-trades careers of the failers of the academic schools is one effective way of maximizing the sum total of our human resources. In wartime this is considered especially important. The picking of no better talent than the job demands is the labor policy often used less consciously by highly specialized industry, or more formally where highly specific motor tests are employed to pick workers for specialized factory work as in the pioneering and effective studies in the Winchester Arms Plant by Link (1919) or the later work by Stead, Shartle, and others (1940).

SYSTEMATIZING PROFILES. It is reasonable to suppose that with only a very few tests available, and these unique in having severally the properties of high validity, approximately zero intercorrelations, and high reliability of each test, we would be able to plot stable personality profiles of persons which would be of utmost value in both guidance and placement. A majority, if not all success criteria, conceivably could be usefully predicted by only a few such tests having these properties.

Having acquired, by several extensive, and expensive, scale-building operations as outlined, about a half dozen such and then designating individual status as tertiles only, we should have 3^6 or 729 different kinds of persons in the world. If the six tests intercorrelate zero, as assumed, all these types would be equally probable and therefore would occur equally frequently so that in a school population of 729,000 children there would be approximately a thousand of each profile. For ease of reference these might be referred to as types 1, 2, 3, . . . 729 easily ascertained from the addends advocated by Toops (1937). Subsequently empirical follow-up tests could be made with the aim of noting in which occupations people of each type succeed best. Thus, eventually, positive guidance, precluding the necessity of selection save as a job-inductive process, would, for the first time, become a reality.

CRITERIA OF PROFICIENCY

Vocational success cannot be easily measured, but it must be measured if we are to weed out the poor tests. We must first obtain each experimental individual's scores in as many aspects of success as are deemed important. If speed and accuracy of typing are considered essential aspects of the success of a stenographer by the employer, then we must measure both. If ability to meet people is highly regarded, then that too must go into the composite. Clearly, also, is the fact that some of the things that we are prone to select may measure almost exactly the same functions as measures already included. To include the same measure twice would overweight that item of success. Any system of combining the component scores of a success criterion is bound to give differential weights to the subparts thereof. It is impossible not to weight the elements if they are added. The system of weighting used, then, is important.

WEIGHTING THE ELEMENTS OF CRITERION. Three current methods of determining the weights of the subparts are in current use: (1) the method of bids devised for the Army S.A.T.C. (1919) makes use of the judgment of experts, who know thoroughly the situation in question, to ascribe up to 100 or 1,000, say, units or bids-of-relative-importance (arbitrary β's) to the several items making up the composite; (2) the method of factor analysis which seeks to ascertain the factors that are involved in the several variables which make up the criterion measures with a view to weighting the several criterion elements in proportion to the extent to which the several variables in turn measure the "central factor," that success which in common is most measured or best measured by the criterion variables collectively; and (3) the method of canonical correlation which seeks to determine which set of test weights may be ascribed to a set of test variables, and which other set of criterion weights may be ascribed to the several criterion elements in order that the former weighted composite shall correlate most highly with, and so predict best, the latter. The first of these methods is perhaps most generally employed. It is clearly described by Toops (1928).

In any case, it is essential in weighting that the different component scores be measured in comparable units before they are multiplied or added. The easiest method of attaining this is to divide the gross scores of each person by the standard deviation of the trait or element in question, thus making the assumption that gross standard scores are comparable units. Ranks, standard scores, and T-scores also are acceptable.

The temptation will be to use, exclusively or preponderantly, foremen's or supervisor's judgments of success. These are often unreliable, a foreman frequently agreeing in his rankings no more than 0.60 with the rating of subforemen in charge of the same men. What is not reliable, of course, cannot be valid, even though the converse statement is not true. It follows that every effort should be made then to get objective measures of quantity and quality of work done under comparable conditions. Equal sales, for example, indicate equal selling abilities only when sales territories are potentially equally productive.

Whole fields of research are largely nullified because of failure to obtain objective and reliable measures for criteria. Bowman (1928) has shown that the validity of teaching aptitude tests is low, has reached and undoubtedly will remain at a low ceiling, until someone is willing to judge teaching-ability by employing a lot more repeated observations of teaching than hitherto have ever been used. The "halo" problem of personality measurement is essentially the issue here.

SALVAGING PREDICTION

If we measure day by day by means of adequate work records, the rate of progress of apprentices in learning a trade, a prognostic test may be unnecessary. This is true because differential rates of growth, really aptitude indices, are ascertainable at an early date. This method is an alternative to building aptitude tests. To administer such requires much more time than a test and may be more costly in wasted effort and money spent on the destined-to-be failures. It does, however, have the merit that the passers are successful on the job itself rather than successful on an only partly valid test which, at best, is only a substitute for the test-by-the-job. A further advantage is that it can be used before adequate aptitude tests are devised. The test of the job also weeds out some of those not emotionally adapted to the job, the potential early-leavers. Tests at present are not very effective in this realm.

It is advisable to be quite imaginative about what constitutes a test worthy of experimental tryout. Many of the activities of the school, shop, playground, home, and vacation resort of the youth, such as are now collected haphazardly, if at all, in cumulative school records, could be utilized if recorded routinely for all and reduced to meaningful scores by means of norms, in making useful predictions of vocational success. Who now knows, for example, what is the significance, in terms of subsequent vocational success, of a boy's being able at the age of ten years to throw a standard baseball 100 feet; what is the significance of his having collected a stamp book of 2,000 stamps at the same age; of having traveled into or through all the states but five at the age of twelve; or of having saved the money for the purchase of a particular pattern of 36 carpenter tools at the same age? Ordinarily we leave the interpretation of such data to the clinician. We need not do this since the item alternative analysis enables us to determine the validity of (*a*) owning a forge, say, scored 1, (*b*) not owning one, scored zero. Accordingly, it follows that whatever the clinician can do by experienced and considered judgment, the statistician, armed with the same data and a definite success criterion, may do both better and far more easily and routinely by his equations and ascertained-by-trial scoring schemes and credits for individual categories of behavior.

NEEDED RESEARCH

Possibly the most useful research finding would be the discovery of a number of measurable factors running through the several thousand human occupations. Thus after adequate analysis, in the nature of testing a random selection of successful workers of each occupation for these factors, one could designate any occupation by a profile number, representative of its trait composition and requirements. Persons having test profiles most nearly resembling the job profile would be those who, other things being equal, have the best chances of success therein. One might draw the profiles of occupations on tracing cloth and superimpose the occupational profile successively on the individual's profile. Other methods are available (Sagebeer, 1938).

Standard tasks, such as the standard airplane flights of the current researches by the Committee on Selection and Training of Aircraft Pilots, subsidized by the National Research Council, will be found useful in many occupations as standard tests of competence on the job. The more complicated the task, the more are these necessary. Few such are now available. The army truckdriver's performance test is one of the most conspicuously successful.

Many questions of statistical theory in criterion construction and even in test theory are still unsolved. One of the chief of these regards the structure of success; whether, briefly, success is best regarded as the *sum* of its weighted subparts or as a function of the *product* of normatized or remodeled units.

Achievement-measuring tests, such as trade tests, promotion tests, graduation tests, and placement tests, as contrasted with prognostic tests, seem to be most lacking in adequate statistical and theoretical background. The excellent work of the army trade tests during World War I, for example, could be immeasurably improved even with present knowledge, and at small cost, in its statistical aspects.

More studies are needed which seek to utilize as predictive variables the qualitative data of personnel forms and cumulative records, such as have been employed by Bittner, Kurtz (1941) at the Life Insurance Sales Research Bureau, and Toops (1935). The utilization of such is beset with problems. First, many of the items are difficult to quantify. Second, some items always are left unanswered by at least a few individuals. The best method of supplying quantitative scores is not surely known. The conviction is growing that the statistician can expertly utilize much data, formerly falling to the clinician, "for proper interpretation."

Statistically identifiable profiles, or patterns of traits of the whole man, not necessarily originated after any particular concept of the gestalt school of psychology, are likely in the future to yield insight into the problems wherein vocational tests are useful. To repeat: "Success is a profile; not a unitary score." Likewise the measure of a man is his test-measured profile, rather than any summated or unitary score.

The traits of one's constant associates, one's wife and one's near fellow-workers and supervisors, for example, are an important part of one's traits, as having causal influence upon one's occupational success or failure. A bad foreman makes of a susceptible employee a bad workman and a low quality of work. A nagging wife is a causal agent in tardiness, absence, and poor work. To utilize such traits: In a basic data table with people's names as ordinates and traits as obscessae, let one vertical section of the table be the workman's own personal traits; a second vertical section, his foreman's traits, perhaps the same traits or those known to be important in foremen; and a third vertical section, his wife's traits, either the same traits or those traits of wives known to be important in their influence on the work-success of husbands.

All of the scores in the body of the table of a given row of the data-table become then the traits of the individual to whom the row applies; and all the traits by columns, may be subjected to multiple regression analysis, factor analysis, and the like. To date these possibilities have been all but completely ignored by statisticians, even though counsellors and psychiatrists have furnished indubitable evidence of their importance. We might factor-analyze them just as we do other test scores. New statistical methods of utilizing them need to be worked out.

In an appreciable number of cases individual success or failure arises as much from the adequacy of "one's opportunity," that is, of one's vocational environment, as from one's personal vocational strengths and weaknesses. The employment of such variables as additional "tests" requires only that one allot a fourth vertical section of the above data-table to the environmental quantitative indices. If not quantitative at the outset they may be quantified. By their employment the probability-of-success tables accordingly could be altered to take account of the fact that one prognostication is appropriate for a given individual (of a certain test-measured profile) on the assumption of his election of one kind of environment while another and different one is appropriate to an alternative election. Tables of this kind should be used in solving the central problem of all psychology, the effective control of human behavior under alternative "incentives."

READING REFERENCES

Bowman, Earle, Unpublished Dissertation for the Ph.D. Degree, Ohio State University Library, 1928.

Chapman, J. Crosby, and D. R. Chapman, *Trade Tests.* New York: Henry Holt and Company, Inc., 1921.

Davis, F. B., *Item-Analysis Data.* Cambridge, Mass.: Graduate School of Education, Harvard University, 1946.

Kitson, Harry D., "Vocational Guidance and the Theory of Probability." *School Review,* 1920, **8**, 143–150.

Link, H. C., *Employment Psychology.* New York: The Macmillan Company, 1919.

Paterson, D. G.; R. M. Elliott; et al., *Minnesota Mechanical Ability Tests.* Minneapolis: University of Minnesota Press, 1930, especially Chap. 9, "The Criterion of Mechanical Ability."

Stead, W. H., et al., *Occupational Counseling Techniques.* New York: American Book Company, 1940, pp. 74, 93–94.

Thurstone, L. L., *The Reliability and Validity of Tests.* Ann Arbor: Edwards Bros., Inc., 1935.

Toops, H. A., *Computational Statistics* (to be published).

———, "The Criterion." *Educational and Psychological Measurement,* 1944, **4**, 271–297.

———, "The Evolution of the Ohio State University Psychological Test." *Ohio College Association Bulletin,* 1945, **113**, 2287–2312.

———, "A Note On Item Selection." *Ohio College Association Bulletin,* 1936, No. 103, 2207–2212.

27. CRITERIA AND VALIDITY

ROBERT J. WHERRY, PH.D.

Professor of Psychology, Ohio State University; Vice-President, Richardson, Bellows, Henry and Company, Inc.; Consultant to Department of the Army and to the Navy Department. Formerly: Assistant Chief in Charge of Research, of the Personnel Research Section, AGO, Department of the Army; Professor of Psychology, University of North Carolina and Cumberland University.

The final and crucial test of any personnel technique is how well it accomplishes the purpose for which it was designed. A test or an application blank or an interview must select those workers who will maximize the profits of a private enterprise or the per unit cost of effective service rendered in a public enterprise. A job analysis method must provide descriptions capable of being used in setting up selection, training and promotion programs. A rating scale must yield a final index which represents the true value of the person or thing rated. A personnel technique which succeeds in accomplishing the purpose for which it was designed is said to possess *validity*. Management has the right to expect that the personnel psychologist can and will demonstrate the validity of the techniques he produces.

The above concept is very simple in theory but exceedingly complex in practice. True validity is frequently difficult if not impossible to establish. Consider more carefully the statement in the paragraph above to the effect that a test must select those workers who will maximize the profits of a private enterprise.

An employee can affect profits in many ways. A fairly complete list of such ways would include:

OUTPUT PER UNIT ITEM

This is an effective element regardless of whether the worker is paid by piece-rate or by straight time. Even the piece-rate worker affects costs for overhead (capital, rent, wages of supervisors, heat, light, etc.) goes on whether he produces a few units or many. His actual pay may constitute a small percentage of the cost of the finished product.

QUALITY OF OUTPUT

The cost of inspection, correction of errors, wasted material, returned goods, loss of customer good will, and failure to meet contracted deadlines may all be increased by a worker who turns out a substandard product.

LOST TIME: ABSENTEEISM, TARDINESS, SICKNESS, DISABLING ACCIDENTS

All of these instances of failure to appear for duty, and of consequent lowered production, force overhead costs to be spread over the fewer items produced and lead to increased costs. Increased sickness and accidents increase the liability of the company for compensation or increase the insurance rates in case the company seeks to protect itself in that fashion.

TURNOVER: QUITTING, FIRING, OR FORCED TRANSFER

The company invests a considerable sum of money in the training of each new employee whether such training is formal or informal. The longer the employee stays on that job the more units of production over which the training cost can be spread. The only exceptions to this rule are those cases in which the employee can be promoted rapidly to some more complex or more responsible job for which the training already given and the experience on the present job represent prerequisite and training respectively. In those cases the loss due to the short time of production on the present job is counterbalanced by the decreased training cost for the higher job.

TRAINING TIME AND PROMOTABILITY

The newly hired employee is always a temporary liability to the company. Even a skilled workman will have to undergo a certain amount of orientation in the policies, rules, and operating techniques of the new company. The usual new employee is seldom skilled and must be trained on the job. The more time required of the supervisor and the longer the period of substandard production, either as to quality or quantity, the greater the cost to the company. Once having mastered the job for which he was hired, he should be capable of promotion after a proper on-the-job training period, to some position of greater responsibility. If the higher positions can be filled internally the training costs will be much lower since the only entirely "green hands" can be instructed in the simpler jobs at the bottom of the promotion ladder where wages and supervision are least expensive.

EMPLOYEE SATISFACTION

A misplaced employee on a job too big or too little, or otherwise unsuited to his interests or capacities is a costly employee. Such an employee is more apt to produce less goods or substandard goods. He is more apt to be absent or feel sick or become an entry in the turnover statistics. If he does not contribute to costs through one of the above channels, it will take more incentives, either direct (sometimes mistakenly called "financial" as if other incentives were not a financial matter as well) or indirect. He is, furthermore, apt to serve as a "focus of infection," spreading his discontent to other employees. The more satisfied employees are originally the less costly it will be to keep them that way.

Any personnel technique in order to be valid, then, must reduce costs in one or more of the above six areas. Since the six areas are not mutually exclusive, but tend to be more or less related to each other, it would probably be safer to say that any successful personnel technique should reduce costs in all these fields or at least that the total net shift in costs, all six areas considered, should be downward. Such sweeping evidence of validity is seldom achieved in actual practice.

Actually no measure of "success" embracing all of the six areas of cost-saving is apt to exist. The technician must instead depend upon such records as exist, or can be constructed, which will reflect some aspect of one of the six areas. Such a record would be called a *criterion;* a group of them would be called *criteria.* A partial list of such *criteria* would include:

1. *Bearing on output per unit time*
    - *a.* Units produced
    - *b.* Number of sales
    - *c.* Items coded
    - *d.* Earnings on a commission basis
    - *e.* Words typed
2. *Bearing on quality of production*
    - *a.* Number of rejects
    - *b.* Cost of spoiled work
    - *c.* Coding or filing errors
    - *d.* Returned goods (sales)
    - *e.* Disgruntled customers (complaints)
3. *Bearing on lost time*
    - *a.* Days present
    - *b.* Number of tardinesses
    - *c.* Days sick
    - *d.* Visits to first aid station
    - *e.* Length of frequency of unauthorized rest pauses
4. *Bearing on turnover*
    - *a.* Length of service
    - *b.* Quits
    - *c.* Discharges
    - *d.* Transfers due to unsatisfactory performance
    - *e.* Transfers at request of employee
5. *Bearing on training time and promotability*
    - *a.* Training time to reach standard production
    - *b.* Cost of material spoiled during training
    - *c.* Rate of advancement
    - *d.* Training courses successfully completed since enrollment (number of jobs in plant for which employee is qualified)
    - *e.* Merit ratings (times recommended for promotion)
6. *Based on employee satisfaction*
    - *a.* Number of grievances registered
    - *b.* Morale survey standing
    - *c.* Visits to plant psychiatrist
    - *d.* Participation in plant athletic contests
    - *e.* Contributions to "suggestion system"

The thirty measures or criteria suggested above are suggestive rather than exhaustive. The alert technician can and frequently must think of many others. Even the partial list suggested, however, will frequently not be obtainable. Their availability presupposes a *record system* unfortunately nonexistent in many plants, often fragmentary, and at best open to suspicion.

As suggested above, the usual practice in industry is to select some one of these criteria on the basis of availability and direct applicability to the technique being validated. Five examples will serve to illustrate the point:

1. A new training program might be validated by showing that the average training time required to reach standard production had been reliably lowered as compared to that required under a method previously used.

2. A selection test for salesmen might be validated by showing that new employees who had made higher scores on the test actually sold more goods than salesmen making lower scores, although both groups had qualified equally on present employment standards.

3. A suggestion system might be validated by showing that there were fewer grievances in the year following its introduction than had occurred in the preceding year in the plant where it was introduced, whereas there was no similar reduction in a plant similar in all respects except that no suggestion system had been introduced.

4. A rating scale for foremen might be validated by showing that ratings obtained by its use would predict the turnover rate of men under the rated foremen.

5. A medical examination might be validated by showing that men with high scores at time of hiring had fewer absences and/or days sick than men with low scores, but alike in all other respects during a 3-month's trial run in one plant.

The above examples are typical of experiments being carried out in industry today and of criteria used in such experiments. They may appear simple, but they are deceptively so. Certain warnings are necessary in connection with the selection of any criterion for a given study. We shall first consider the criterion cited in example 2, and then present a check list to be used in the selection of any criterion.

In example 2 the criterion was said to be "amount of goods sold." Such a statement is an over-simplification of the problem. Otis and Cooper (Stead, Shartle, 1940) describe at length the various data collected by the Worker Analysis Section of the United States Employment Service in establishing a criterion for department store salespeople. The uninitiated technician might feel that

1. *Gross sales:* total face value of salesslips written, would be an adequate criterion, but Otis lists the following additional data collected:
2. *Number of sales:* the total number of customers to whom such merchandise was sold.
3. *Actual credits:* value of merchandise returned accompanied by salesslip bearing the clerk's name or number.
4. *Returns to stock:* value of merchandise not sent out of store because of poor credit standing of the customer and goods sent C.O.D. but not accepted.
5. *Department credit:* value of merchandise returned without a salesslip, apportioned to each clerk in a given department on the basis of number of sales.
6. *Total net sales:* gross sales minus actual credits minus returns to stock minus department credits.
7. *Average per sale:* total net sales divided by number of sales.
8. *Daily sales quota:* daily salary of clerk divided by average daily selling cost of the department over the last five years.
9. *Days selling:* number of days worked.
10. *Actual quota:* daily sales quota multiplied by days selling.
11. *Ratio: net sales to quota:* total net sales divided by actual quota.
12. *Salary:* amount paid to salesperson as salary, bonuses for exceeding quota, and extra bonuses or "plus money" paid for selling slow-moving items.
13. *Cost per cent:* salary divided by total net sales.
14. *Over-all evaluation:* composite rating of each clerk by personnel manager or supervisor; in which of five classifications of salespersons.
15. *Paired comparison rating:* converted rank scores based upon number of times each salesclerk was voted best when compared with every other salesclerk in the department.

All of these measures are clearly not of equal importance nor do they measure the same things. Cooper reports the correlations between a composite of items 14 and 15 with certain of the other items, as shown in Table 1. On the basis of these correlation coefficients we might conclude that *Net sales per day* was the best of the five objective criteria, a finding not at all out of line with the reasoning back of the original collecting of the data. Item 1, corresponding to the criterion suggested in the hypothetical experiment in example 2, when corrected for such factors as items 2, 3, 4, and 5, yielded item 6 which in turn needed to be corrected by item 9 to make the measure comparable for all individuals.

*Table 1: Coefficients of Correlation between Objective Criterion Variates and a Composite Subjective Criterion for Department Store Salespersons**

Objective Criterion Variate		Coefficient of Correlation with a Subjective Criterion (14 and 15)
2 ÷ 9	Number of sales per day	0.05
6 ÷ 9	Net sales per day	0.49
8	Daily sales quota	0.23
7	Average per sale	0.34
13	Cost per cent	−0.15

* Adapted from Cooper.

Item 8, daily sales quota, with considerably lower agreement with judged ability, represents another type of "poor" criterion frequently chosen by the unwary personnel research man. The largest single element in the denominator of that ratio was "salary" which is frequently based upon "seniority" rather than upon "merit." The untrained personnel man, frequently awed by complicated processes which he doesn't understand, might have chosen item 13, cost per cent, as his criterion. The fact that this criterion was negatively related to supervisor judgment of the salesclerk should not be surprising and is easily explained. Salary (item 12) *should* be a perfect reflection of productivity (item 6). Since each should measure the competence of the individual, the dividing of one by the other to obtain "cost per cent," item 13, would then partial out all ability. Actually salary tends to lag behind productivity since the lowest producers are often grossly overpaid due to minimum wage laws and inertia and legal restrictions which mitigate against separation of such employees. The high producers therefore tended to have lower ratios and the ratios (item 13) tended to correlate negatively with the ratings (item 15 and 16).

It should be evident from the foregoing that the selection of a satisfactory criterion is not an easy matter. Bellows (1941) has suggested several checks for estimating the goodness of a criterion which should prove of value in making such a selection: *Is the proposed criterion statistically reliable?* In the study of salesclerks by the USES cited above, Cooper found that the reliabilities of items 14 and 15 were 0.75 and 0.71 respectively. The more objective measures showed considerably higher reliability, with "net sales per day," item 6 divided by item 9, being 0.94. While reliability alone is not a sufficient reason for selecting a criterion, it should be remembered that if the criterion is not reliable it will be of little value as a test of the personnel technique being evaluated. *Does the proposed criterion correlate with other* *criteria?* The value of this device is again shown by the USES salesclerk study. In Table 1 we saw that "net sales per day" correlated most highly with the subjective ratings, whereas "cost per cent" actually correlated slightly negatively. While the subjective ratings were finally discarded as of lower reliability than the objective measures, they did serve to suggest that of the objective measures "net sales per day" was a much better selection than "cost per cent."

Is the proposed criterion acceptable to the sponsor of the study? When six experts chosen from section managers and the personnel department rated the various criteria for importance, the "net sales per day" variate was considered the most important and most representative of job performance. While this standard should probably never be used alone, it should be taken into account. When, as in this case, it *supports* other data, it is a valuable addition. A criterion in which the sponsor has confidence will help in the final "selling" of the tested technique if it proves valid, and a new technique will be of no value unless it is adopted and put into operation.

Is the proposed criterion acceptable to the job analyst? We might rephrase this question to read, "Does the proposed criterion measure those aspects of work performance given heavy weights in the job analysis?" Actually this question is the best starting point for collecting a list of potential criteria and is probably safer than the alternative procedure of asking what data are already available. The fact that, say, "number of clerical errors on sales slips" had already been tabulated, whereas "net sales per day" would require considerable calculation would never justify the use of the less important of these indexes of salesmanship as *the* measure of success.

Is the proposed criterion predictable? If several predictors, covering various aspects of job competence, and several criteria, covering various aspects of job performance, are available, then that criterion which correlates most highly with the most potential predictors would tend to be selected. At the beginning of this discussion the fact was stressed that job success is a complex affair. To permit all important predictors to get into the final selection battery, the criterion must contain measures of all the important facets of job success. This check is particularly important when the validation is of a battery or program rather than of a single test or device.

Has the proposed criterion when used in the past lead to selection of instruments which have produced practical beneficial results? If previous techniques, either constructed or selected on the basis of this criterion, are judged or, better yet, can be demonstrated to have resulted in increased efficiency of operation, that would be a strong argument in favor of continued use of that criterion, *other things being equal.* If other things are not equal, this check should not be allowed to stand in the way of progress and even better criteria. Dr. Edward E. Cureton, in a private conversation, suggested an interesting ramification of this concept. He proposed that the personnel department keep records of all potential criteria over a period of several years and correlate these with actual earnings of the company (or department) to see which were really important determiners of costs and profits.

The Bellows article also discusses at length three types of criterion contamination. Even though a criterion meets all six of the standards set forth above, it may still be made almost worthless by any of these kinds of contamination. The research worker must always take care to eliminate any possibility of their occurrence.

The first kind of contamination discussed is that arising from illicit knowledge of the predictors by persons responsible for the criterion. Assume, for example, that an experiment is conducted to discover the validity of an intelligence test for the selection of foremen and the criterion of success as a foreman is promotion to assistant superintendent. Suppose that the intelligence test scores are made known to the general manager and that he actually uses them as the basis for promoting foremen to assistant superintendent. Or, suppose, merely, that superintendents are made aware of the scores made by the foremen and devote training time and help primarily to those with low scores while those with high scores are left to shift more or less for themselves. In the first case the discovered relationship between test score and success would be very high and the experiment would seem to indicate high validity. In the second case the extra aid given low score people might counterbalance the lower ability so little or no relationship would exist between score and foreman success, and the experiment would indicate zero validity. Actually, in neither case would the experiment have been a legitimate test of the predictor. The results would have been dictated by illegitimate use of the predictor scores rather than by actual difference in intelligence. We conclude then that *in experimental situations scores on predictor items must be kept secret until the end of the experiment if contamination by illicit use of such information is to be kept out of the criterion.*

The second type of criterion contamination discussed by Bellows is that due to *artificial limitation of production.* Actual output on the job is always a joint function of the worker's traits, capacities, abilities, etc. and of the tools, equipment, work surroundings, condition of raw materials, etc. If these latter differ significantly from worker to worker, they may completely mask differences due to worker qualities. Failure to equate or control working conditions is just as fatal to industrial experiments as is lack of control in any other experimental situation. It is the inability to control such factors that has lead to so many negative (low or zero validity) reports of industrial selection experiments when the criteria have consisted of actual production records. It is this type of contamination that lead us to say earlier that all available criteria must be regarded with suspicion.

Frequently, these two types of contamination may be compounded. A supervisor who believes strongly in one of the predictors may give favored spots, better material, better machines, and other advantages to persons known to have high scores on that particular variable. In such cases the resulting validity coefficient might be quite high and yet be almost completely meaningless and unreliable in the sense that for future samples with other supervisors the results would turn out quite differently.

It is this second source of contamination due to artificial limitations of production which has lead many research workers to abandon production data as a source of criteria. In its place they recommend securing what is called a *work sample.* Differences in material, complexity of process, efficiency of machine, lighting, availability of raw materials, and the like, can be kept constant if all workers are put through an *identical* replica of the production process. The work sample must of course meet the six checks described above for any criterion. In the case of a complex job it may be necessary to prepare and standardize a series of such work sample situations in order to assure representation of all major types of duty shown by the job analysis. The important consideration is to assure that all contamination due to artificial limitations of production are removed by having all workers qualify under identical conditions.

One very interesting work sample procedure was worked out by Viteles (1945) in connection with licensing and qualifying examinations under the Civil Aeronautics Administration. Inspectors frequently tested candidates under widely diverse conditions and with ingenious variation in requirements. To secure a criterion free of such contamination Viteles developed what he called *standard flights.* These standard flights "provided, for use in the air, standard patterns of maneuvers analogous to the standard series of driving operations used both by civilian agencies and by the military sources in measuring proficiency in the operation of a motor vehicle." It is probable that industrial research will make wider and wider use of this *work sample* approach as it is realized how widely different can be such simple operations as typing *a* letter, punching *a* card, making *a* tire, or fixing *a* leak. Standardization of criteria may in time come to be recognized as equally important to standardizing a test or other predictor variables.

The third type of contamination suggested by Bellows was that due to *differential influence of experience.* This type of contamination can be extended to refer to sex, age, and possibly other personal, as contrasted to job environment factors as well as to job experience or job training which was inferred in Bellows' discussion. That job training is an important factor was shown by some material concerning lampshade workers reported by Otis and presented here in Table 2.

Table 2: Distribution of 18 Inexperienced and 19 Experienced Lampshade Sewers, According to Their Average Units of Production per Hour

Average Units of Production	Experienced Workers	Inexperienced Workers
49–41	6	
45–48	1	
41–44	5	
37–40	6	
33–36	1	3
29–32		3
25–28		2
21–24		3
17–20		4
13–16		3

Unless statistical correction were made to production figures on the basis of experience, no test of aptitude would be apt to correlate with such a contaminated criterion.

This same type of contamination may appear on the prediction as well as on the criterion side of the correlation. Male and female workers may perform equally well a specific skilled task on which they have had sufficient experience, but scores on an aptitude test such as the Bennett Mechanical Comprehension Test, would be much lower for women than for men workers. A validity coefficient relating prehiring test scores and later performance would be entirely too low unless a differential credit for sex was achieved by using separate test norms for each sex.

Another way of removing this last type of contamination would be to stress *homogeneity* of research population. This technique of restricting the experimental population to a single level of age, sex, experience, and similar factors is a procedure carried over from the classical research field. When populations are large enough to permit such rigid selection of cases this concept offers a satisfactory alternative to statistical correction. In industrial situations large populations are the exception rather than the rule, and it is frequently necessary to include all available workers in order to obtain a sample of adequate size. Otis points out that, while statistical correction produces some error in the data, this error is usually smaller than that produced by the unequal influence of experience on the scores. Another point in favor of statistical correction is that if homogeneity is used it will be necessary to conduct many studies (one for each sex, each age level, each experience level, etc.) and to evaluate these combined findings in order to gain an over-all estimate of validity.

A third way of removing this type of contamination is to include the offending variables (age, sex, etc.) as separate predictors in the experimental battery and properly evaluate their corrections by multiple correlation techniques. While not requiring populations as large as would the homogeneity concept, this approach is still not too satisfactory if the experimental population is small and the prospective battery already large since the restriction on degrees of freedom caused by the added variables may render the results unduly subject to chance errors.

When the research worker has tested each proposed criterion by means of the 6-point check list and has assured maximal freedom from the three types of contamination, he may have left only a single outstanding criterion. If this is the case, he would adopt that criterion which was most important or most outstanding and proceed with his study. Frequently, however, there will remain several criteria of almost equal merit. This situation presents the technician with two major alternatives.

1. Two or more of the measures may be retained and an attempt made to predict them independently.

2. The various measures can be combined to give a single index of occupational proficiency. Which of these two alternatives is best to follow can be best decided by a consideration of the correlations among the acceptable criteria, if these are known or obtainable.

If the various criteria are actually independent, as shown by zero or near-zero intercorrelations, any composite function will be more or less meaningless. If, on the other hand, the criteria show uniform and moderate to high intercorrelations, some form of composite is indicated.

Horst (1936) has given a method of weighting criteria so as to maximize the dispersion of the composite scores. Edgerton and Kolbe (1936) worked at the problem from the point of view of minimizing the dispersion of individual scores before combining, which resulted in the same set of combining weights. These results were in turn shown to be the same as the weights of the first principle component as derived by Hotelling (1933). Each of the above approaches requires extensive and complicated mathematical operations, but Wherry (1940) has developed an approximation method which greatly simplifies the finding of the necessary weights for combining the variables.

In case the correlations among the criterion variables are varied, ranging from high to low, the best procedure would be to do either a Hotelling principle component or a Thurstone centroid factor analysis. In either case the mathematical factor structure should be rotated into a meaningful pattern. Each set of meaningful factor loadings would in turn be used as "criterion coefficients" along with the intercorrelation table to secure a set of combining weights, by multiple correlation, to best predict that factor. A simpler but much less reliable approach would be to select that criterion variable with the highest factor loading from each factor as most representative of the cluster in question.

In case the intercorrelations are unknown, the technician might employ a method suggested by Toops (1928), the *Estimated Beta* method. In this method several experts are selected and then asked to assign some fixed total of points (100 or 1,000) to the various predictors. The average assigned weight for each variable is then determined. It is assumed that these weights apply to standard scores of the criterion variates and are divided by the standard deviations of the respective criteria before being used as given score weights.

If the investigator wishes to make the highly unlikely assumption of equal intercorrelation he could follow the frequently used practice of reducing all criterion raw scores to *standard* scores, and then simply proceed to add them up, thus in effect using weights of unity.

While no one of these methods appeal to the research worker, they represent the best opinion of what to do when in the embarrassing position of having more than one equally satisfactory criterion.

READING REFERENCES

Bellows, R. M., "Procedures for Evaluating Vocational Criteria." *Journal of Applied Psychology,* 1941, **25,** 499–513.

Edgerton, H. A., and L. E. Kolbe, "The Method of Minimum Variation for the Combination of Criteria." *Psychometrika,* 1936, **1,** 183–187.

Horst, Paul, "Obtaining a Composite Measure from a Number of Different Measures of the Same Attribute." *Psychometrika,* 1936, **1,** 53–60.

Hotelling, Harold, *Analysis of a Complex of Statistical Variables into Principle Components.* Baltimore: Warwick & York, Inc., 1933.

Stead, W. H., et al., *Occupational Counseling Techniques.* New York: American Book Company, 1940, Chap. 4, "Rating Forms"; Chap 5, "The Criterion."

Toops, H. A., "The Selection of Graduate Assistants." *Personnel Journal,* 1928, **6,** 457–472.

Viteles, M. S., "The Aircraft Pilot: Five Years of Research; A Summary of Outcomes." *Psychological Bulletin,* 1945, **42,** 489–526.

Wherry, R. J., "An Approximation Method for Obtaining a Maximized Multiple Criterion." *Psychometrika,* 1940, **5,** 109–115.

28. ITEM ANALYSIS

ROBERT J. WHERRY, PH.D.

Professor of Psychology, Ohio State University; Vice-President, Richardson, Bellows, Henry and Company, Inc.; Consultant to Department of the Army and to the Navy Department. Formerly: Assistant Chief in Charge of Research, of the Personnel Research Section, AGO Department of the Army; Professor of Psychology, University of North Carolina and Cumberland University.

Not every question or incomplete statement is a test item, nor does every group of test items constitute a test. Further, a test which proved quite efficient in selecting clerical workers in Milwaukee in 1946 may not, and probably will not, constitute a satisfactory test for other clerical workers in Detroit in 1951. Many testing programs are abandoned each year because the sponsors either do not know or do not understand these simple truths.

The selection of an item from among a group of questions, the grouping of items together to form an effective test, or the "retailoring" of an existing test to fit a present need is accomplished through what the psychologist calls item analysis.

Item analysis is no substitute for sound item writing or wise preliminary selection of testing material. The role of a thorough situational and job analysis as a preliminary step to any collecting of questions or tests cannot be too strongly stressed. Item analysis cannot make a silk purse from a sow's ear; it can only tell which is which.

On the other hand, the use of trained job analysts and skilled item writers will be no substitute for proper item analysis. Even the batting champion of the American League seldom achieves an average as high as 0.400 for a given season. Similarly a skilled item writer, obeying all the rules, will write more poor items than good ones, even when he is thoroughly familiar with the area in which he is working. Expert writing and expert review can increase the frequency of satisfactory items, but only sound item analysis can discriminate the good items from the bad.

Item analysis may be used to select the best items from a larger number in order to achieve maximal results within the testing time allocated. Just as motion analysis is used to eliminate needless or harmful motions from production methods, so item analysis may be used to eliminate needless or harmful items from a test. A test pruned in this fashion will select employees better and faster.

Item analysis can also be used to make the same test achieve different or multiple results. An automobile would be of value even if it moved only in a forward direction in low gear, but the reverse direction and variable forward speeds achieved through complicated gearing make it much more efficient. Just so a test with a single scoring key may be valuable, but the same test with several scoring keys, each designed to accomplish a special selection job, will make it much more valuable. It is through item analysis that such multiple keys are constructed and kept up to date. The multiple scoring keys of many interest or personality inventories are cases in point.

Item analysis may be conducted in various ways depending upon which of the above purposes is to be accomplished. It may center around the difficulty of the test questions. It may be based upon agreement between the action of the single item and that of a large group of items or of the entire test. Or it may be based upon the agreement between the single item and an external yardstick. Each type of analysis has utility, but the various approaches are not interchangeable. The skilled test technician must be able to use all approaches and know when to use each one.

The use of any item analysis technique depends upon the availability of a trial administration of the proposed test material. The larger the sample population available for such trial the better. The persons used in this trial should be representative of the persons to whom the finished test is to be administered. For the third type of analysis, the comparison of the item with an external yardstick, certain job performance or other descriptive information must be available for each person.

DIFFICULTY ANALYSIS

The simplest type of item analysis is that based upon item difficulty, i.e., upon the proportion of the population passing or failing each item. If an item is to differentiate people, people must differ in the way they respond to the item. If an item is so easy that everyone gets it right, it is a worthless item. If it is so hard that everyone misses it, it is also worthless. By worthless we mean merely that neither type of item could tell us which of a group of 100 applicants was better or poorer than the other.

Item difficulty is computed by dividing the number of people who pass the item, called R for right answers, by the number of people, called N, who were exposed to the item. This ratio expresses the probability, p, that any given applicant will pass the item. Thus we can write the very simple equation $p = R/N$.

For an item which everyone passed, p would equal 1.00; while p would equal 0.00 for an item which everybody missed. Since the item writer generally tries to get a wide spread of item difficulty he will usually write some items with these extreme values. Since these items only cause extra printing, administration, and scoring costs, they should for the most part be eliminated from the test. There are, however, uses for a few such items.

A few of the very easy items, with $p = 1.00$, can be utilized as practice items so that everyone will understand exactly what is to be done. In addition a few such items are frequently placed at the beginning of the actual test to motivate test takers of low caliber. The presence of a few such items will also assure the absence of a floor to the test so that if in future populations the ability should be slightly lower than that of the trial population, it will still be possible to differentiate between the very poorest members of the group.

Similarly the majority of very difficult items, with $p = 0.00$, should be eliminated. However, a few such items should usually be retained to provide ceiling for the test. Such ceiling is provided in order that the test will still differentiate the very best applicant in some future population with some member of even greater ability than anyone in the sample population.

It was stated above that item writers strive to attain a wide spread in item difficulty. This is true because they know that items of varying difficulty each accomplish a different task. Assume for the moment that all items work perfectly, i.e., that if p people pass the item these are the p best persons in the group. A single item with a p of 0.50 would then merely sort the population into an upper half, those who passed the item, and a lower half, those who failed the item. The addition of other items with p equal to 0.50 would give us no extra information since each in turn would merely verify our previous grouping. An additional item with p equal to 0.25, however, would split the upper half of the population into upper and second quarters. Similarly, an additional item with p equal to 0.75 would split the lower half into third and fourth quarters although it would give us no new information about the upper half of the population. Finally, in order to make differentiations all along the line we would want to have items of all possible levels of difficulty.

Of course we have exaggerated the efficiency of items in the above discussion since many people will get the correct answer solely by chance and since items seldom measure exactly the same thing. Only in very rare cases would two items with p equal to 0.75 be passed by exactly the same 75 per cent of the population. There would in most cases, however, be considerable overlap in the list of persons passing the two items. In general the discussion in the previous paragraph applies well enough to lay down three principles for the selection of items on the basis of difficulty.

1. To divide the population into many levels and have as few tied scores as possible, use items from all possible levels of difficulty.

2. To eliminate a few people of low ability at the bottom of a group of applicants, keep predominantly items with high p values.

3. To skim off only a few of the better applicants from the top of a large group, keep predominantly hard items with low p values.

If the type and number of applicants is known and stable it is possible to use principle 2 or 3 to shorten the testing program. If the type and number of applicants is unknown or unstable it will be safer to base the test construction upon principle 1. In a tight labor market with few applicants one would probably use principle 2, while in an easy labor market with a large surplus of applicants tests could be built around principle 3.

A consideration of the shape of the population distribution will give further insight as to the most desirable distribution of item difficulties. In an unselected population the best bet is that the population will be normally distributed, that is, will pile up in the middle and taper off toward the ends of the distribution. This means that there will be more people to differentiate in the neighborhood of p equal to 0.50. Since items are not perfect, there should be more items with p equal to 0.50 with successively fewer items as one approaches p equal to 0.00 or 1.00. In selected populations the distribution of ability will not always be normal, but we can state a fourth general principle.

4. The distribution of item difficulties should be similar to that of population ability to best differentiate all persons in the group.

One final use of difficulty analysis is the arrangement of items in the completed test. In general, items should be presented from easiest ($p = 1.00$) to hardest ($p = 0.00$). This is especially important in speed-tests or tests with time limits. A hard item placed near the beginning of the test might take the entire testing time for persons of relatively low ability, and it would be impossible to say whether they might have answered 1 or 5 or 20 of the easier items appearing later in the test. If, however, the items are arranged in descending order of p value (ascending order of difficulty), each person will have achieved his proper level before he is stumped by any particular item. Research studies have shown that arrangement of items in this way improves both the reliability and validity of test results.

ITEM-TEST ANALYSIS

A second type of item analysis is that based upon the relationship of the item response to total test score. All papers are scored for total number of correct responses. The papers are then sorted into upper and lower halves, upper and lower quarters, or some other high and low groups. For each item the number of correct responses is then determined for the upper and lower groups separately. If the item belongs in the test, that is, if it is measuring the same thing as the test, the number of correct responses in the upper group will be much larger than the number of correct responses in the lower group. The larger this difference, other things being equal, the better the item. The actual worth of the item would be determined by some variety of correlation coefficient, or some approximation to a correlation coefficient. The test analyst should be familiar with several different methods to most effectively handle any given type of data. Among the easiest methods to use are the abacs or computing charts published by Mosier and McQuitty (1939). There are many short-cut approximation methods in the literature and these are frequently satisfactory. The computation of the actual correlation coefficients, however, will enable the technician, not only to arrange his items in order of merit, but also to tell whether the result achieved by any given item is better than could have been expected by chance.

Items with high coefficients should be retained for use. Items with negative or chance positive coefficients should be eliminated. The elimination of such items not only shortens the testing time but actually improves the effectiveness of the test since such items add only error or chance to the test scores of the individuals.

After this first purification has taken place the entire group of papers should be rescored on the basis of the retained items and the entire process should be repeated. This iterative process should be carried on until the last total test score used was based almost entirely upon finally retained items.

A word of caution should be inserted at this point. The above method of analysis is applicable only to aptitude or proficiency tests where a single ability or factor is involved. If the test is of omnibus form, i.e., measures several things at the same time, the procedure will result in keeping only the most highly represented types of material. In interest or personality inventories, or omnibus proficiency or aptitude tests, a poorly trained analyst might throw away as useless many valuable items. For example, if an intelligence test (usually omnibus in form) contained 60 verbal items, 30 number items, and 10 space items, all of which were perfectly reliable, the verbal items would correlate 0.77, the number items would correlate 0.55, and the space items would correlate 0.31 with total test score. This might result in the misguided discarding of the 10 space items. True, one would then have a better verbal and a better number test but a much poorer measure of intelligence, especially if the test were to be used to hire file clerks, machinists, or other employees who need spatial ability. For such cases Wherry and Gaylord (1946) have presented a method for selecting successively several subsets of items from such a complex test.

ITEM-CRITERION ANALYSIS

A third type of item analysis is based upon the relationship of item response to some external yardstick or criterion. The test population must again be divided into upper and lower groups as in the previous analysis, but this division is made upon a criterion of success not directly based upon test performance. The value of the analysis will depend upon the goodness of the criterion used. Extreme care, as much as that used in the construction or preliminary selection of the test material, should go into the selection of the criterion to be used. It must not only be a reliable, accurate measure of success on the job, but must embrace, and preferably be limited to, success in that particular aspect of the job which the items to be analyzed were designed to measure. In case of an omnibus test covering several abilities or qualities required for the job, it may be necessary to have a series of criteria relating to performance on various aspects of the job.

The actual testing of the items, once the criterion has been settled upon and the high and low performance groups established, proceeds in the same manner and utilizes the same techniques as does item-test analysis. The resulting item validities are interpreted in the same manner, but are restricted in applicability to the specific job upon which the performance records of the subject were based. If the requirements of the job are changed at some future date the items may no longer be valid for the changed job.

CROSS VALIDATION

One important caution applying to all item analysis but particularly to item-criterion analysis, must be stressed. Impurity of test or criterion or deviation of the sampled population from total population frequently results in wide fluctuation of item coefficients from sample to sample. A test composed of items with such fluctuating coefficients is apt to yield disappointing results upon reapplication. The solution to this problem is multiple cross-validation.

In multiple cross-validation the total population is broken into several random or equivalent groups. Only those items are chosen for final retention which show little fluctuation in magnitude and no reversal of sign or direction of validity, and for which the over-all or total discrimination meets some minimum standard. It has been stated by some experts that personality tests frequently go through as many as 10 revisions before the items steady down. This means that with the usual approach of a single trial population it might take several years to perfect a test.

The idea of immediate multiple item analysis was adopted by the Personnel Research Section of the Adjutant General's Office with excellent results. In selecting items for a Biographical Information Blank, nine independent samples of from 50 to 200 cases each were collected and analyzed separately. When only items which had maintained satisfactory indices on all nine samples were selected to form the original test, all items were found to hold up on two new samples collected later over a period of 6 months. This seems to indicate the inadvisability of the small type of pilot study now frequently used by industrial research departments if larger multiple cross-validation studies are possible.

COMBINED TEST-CRITERION ITEM ANALYSIS

Maximal benefits can be derived by combining the item-test and item-criterion types of analysis. Such combination is based upon the theory underlying multiple correlation. No attempt will be made to give an exhaustive explanation of this theory here, but two generalizations can be made.

1. Other things being equal, the higher the item-criterion index of the item to be added the better the item.

2. Other things being equal, the lower the relation of the item to previously added items, the better (a new area of content or a new difficulty level will add more to the discrimination of the new total scale). While generalization 1 can be applied immediately to the item-criterion coefficients, the information concerning item to item relationship can only be obtained at great expense. As a substitute, however, we can use the item-test coefficient (how the item agrees with all other items) in place of the *item-previously used items* coefficients. Two systems based upon the substitution appear in the literature.

Horst (1936) worked out one of the two systems. Item numbers are plotted in a double entry table with the item-criterion values plotted horizontally and item-test values plotted vertically with the axes intersecting at the 0.00 point on each scale. A movable pointer, like the hour hand of a clock is fastened with its base at the intersection of the axes and pointing toward 3 o'clock. This pointer is then moved anti-clockwise through the first quadrant until enough items to fill the allotted testing time are found beneath the moving indicator (all items with negative validity are rejected, all items with positive validity and negative test correlation are accepted). This tends to maximize the item-criterion values while minimizing the item-test coefficients. It is a very useful tool when selection of items must be accomplished in a hurry. If time allows, the item-test coefficients should be recomputed (the item-criterion values will not be affected) on the basis of the items actually selected, and the process repeated until few new items are selected. This will, if thus used iteratively, give a very exact approximation to the best multiple correlation.

Another system utilizing the two indices was worked out by Flanagan (1936). His process is based upon item-weighting rather than upon item-selection. All items are given an initial weight of plus one, and the item-criterion and item-test coefficients are computed. The item weights are then changed as follows:

1. If the item-criterion correlation exceeds the item-test coefficient, *increase* the weight by one point.

2. If the item-test correlation exceeds the item-criterion coefficient, *decrease* the weight by one point.

3. Recompute total test scores on the basis of the new weights, recompute the item-test correlation coefficients and repeat steps 1 and 2. This procedure should be repeated until no or few further changes take place. In general, no difference is recognized unless the two coefficients differ by at least 0.10. A few iterations will give a fairly exact approximation to the true multiple. Items with zero weights can then be omitted or simply scored zero as desired.

The Horst method is best when the length of the test is predetermined by testing time available or when it is desired to use only unit weights. The Flanagan method is best if testing time is not restricted or if variable weights are allowable. In either case, if items are omitted, the shortened test should be revalidated upon a new population to discern the effects of changed associations arising from surrounding items. Cross-validation is important in any case, however, since like multiple correlation itself, the two methods tend to utilize chance as well as true relationships. If the test analyst is willing to compute several more coefficients he can get an exact solution directly without iteration by either the Horst successive residual method (1934), the Toops L-Method (1941), or the Wherry-Gaylord L-Method (1946). The second of these methods permits only unit weights, while the first and last methods permit either unit or variable weights as desired.

ALTERNATIVE ANALYSIS

In the discussion thus far we have been assuming that the analyst was dealing with items in which the correct answers are known in advance. The test analyst refers to what we have been discussing as *rights analysis*. In certain areas, particularly in those dealing with interests, personality traits, or rating phrases the correct answers are frequently unknown, in which case the technician must first resort to *alternative analysis*.

In alternative analysis the high and low groups must be determined on the basis of an external criterion, since test score is unknown until after right and wrong responses are determined. Then, instead of recording number of rights for each group, the number of responses to each alternative for each group is recorded. The number of such responses in the high group minus the number of responses in the low group, divided by the number of people choosing the alternative can be used as a *discrimination index* for that alternative. A plot of these alternative indices from highest to lowest will indicate where the sharpest break or breaks occur in the item. If a single sharp break occurs, all alternatives above this point would be considered as Right responses. If two such sharp breaks occur, items above the highest break would be scored as *right* and those below the lowest break as *wrong,* alternatives between the two breaks would not be scored. If all positive scoring is desired all items below the lowest break can be scored zero, and one extra point can be assigned for each successive sharp break above which a given alternative lies. A refinement of these weights could be achieved by using them as a starting point in the Flanagan process described above. Multiple cross-validation is extremely desirable for alternative analyses.

Alternative analysis is also frequently employed in item-construction as well as in item-scoring. In aptitude testing or achievement testing the wrong alternatives should all possess negative validity that is, have negative discrimination indices, and should all have approximately equal pulling power, that is, equal choice by those not giving the correct answer. Horst (1933) has written concerning the theoretical background of this last requirement, and Toops and Adkins (1937) have used the procedure in constructing items for the Ohio State Psychological Examination. Items are administered to the trial population as completion (fill-in) questions. Each answer given by a test subject is treated as an alternative. The best and most positively valid right response and negatively valid and equally frequent wrong responses are then chosen to give sufficient alternatives for a multiple choice item.

After the items have been properly constructed and the proper scoring key has been established by alternative analysis, the test technician can then submit the items to the same type of difficulty, item-test and item-criterion analysis described in earlier sections.

READING REFERENCES

Adkins, D. C., and H. A. Toops, "Simplified Formulas for Item Selection and Construction." *Psychometrika,* **2,** 165–171.

Flanagan, J. C., "A Short Method for Selecting the Best Combination of Test Items for a Particular Purpose." *Psychological Bulletin,* 1936, **33,** 603–604.

Horst, Paul, "The Difficulty of a Multiple Choice Test Item." *Journal of Educational Psychology,* 1933, **24,** 229–232.

———, "Item Analysis by the Method of Successive Residuals." *Journal of Experimental Education,* 1934, **2,** 254–263.

———, "Item Selection by Means of a Maximizing Function." *Psychometrika,* **2,** 227–244.

Mosier, C. I., and J. V. McQuitty, "Methods of Item Validation and abacs for Item Test Correlation and Critical Ratio of Upper-Lower Difference." *Psychometrika,* 1940, **5,** 57–65.

Toops, H. A., "The L-Method." *Psychometrika,* 1941, **6,** 263–264.

Wherry, R. J., and R. H. Gaylord, "The Concept of Item Readability in Relation to Factor Pattern." *Psychometrika,* 1943, **8,** 247–264.

———, "Test Selection with Integral Gross Score Weights." *Psychometrika,* 1946, **11,** 173–183.

29. PERSONALITY TESTS AND PREDICTION

EDWARD A. RUNDQUIST, PH.D.

Assistant Director of Personnel Research for the Owens-Illinois Glass Company; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology; Consultant on Measurement of Personality, Richardson, Bellows, Henry & Company. Formerly: Member of the Staff, Personnel Research Section, AGO, Department of the Army; Psychologist for the Minneapolis and Cincinnati Public School systems; Institute of Child Welfare, University of Minnesota.

To personality characteristics is usually assigned a major role in occupational success. Estimates of them are relied on heavily in selection of supervisors, executives, salesmen, and even unskilled and semiskilled workers (Achilles, 1946). A view widely held is that for supervisory positions personality characteristics are more important than skill or technical know-how. The importance conceded to them and the interest inherent in the subject has stimulated a wide variety and volume of research on tests of the direct-question type (such tests as those of general personality tendencies, attitudes, social and ethical values, interests); on tests utilizing the projective principle (ink blots, incomplete sentences, interpretation of pictures); and on efforts to measure personality trends indirectly by means of tests which appear to be tests of perception, learning, or aptitude. This work has defined problems and increased our general understanding of personality adjustment, but the contribution to the prediction of specific occupational success has been negligible. Tests of interest have proved of some use in guidance but not as effective in prediction of degree of success in a given field. Tests of the general adjustment variety were found to be of some usefulness in neuropsychiatric selecting in the Armed Forces (Hunt and Stevenson, 1946), i.e., for identification of those so maladjusted that their likelihood of success on almost any job is small. While this is important, it is an approach to the negative aspect of the problem of prediction. The long-range goal is to discover methods of predicting success in a particular kind of work.

As a result of the lack of objective measures of known significance for specific jobs, chief reliance has been placed on the interview in the estimation of personality characteristics. Tests of personality, if used at all, are evaluated subjectively or used as points of departure for interviewing.

Under certain conditions the interview has been shown to have some promise as a selection device (McMurray, 1947; AGO, Report 704), but as usually applied in a brief and unsystematic way, it is doubtful if it is more valid than the usual personality scale. Three features of the interview—its susceptibility to the operation of bias, the difficulties in finding a consistent frame of reference for the expression of its results, and its time-consuming nature—make the search for more objective techniques for the measurement of personality highly desirable.

It is the purpose of this section to consider briefly the reasons for the general failure of personality testing, and to describe the two techniques of personality test construction which show particular promise in predicting occupational success. One of these techniques is merely the full exploitation of personal history and background information and may be regarded as an extension of the application blank; the other is a new technique for the construction of personality items: it is termed the forced-choice technique.

REASONS FOR LACK OF SUCCESS

Of the many reasons for the failure of traditional type personality tests in occupational guidance and selection, three deserve special mention: (1) the ease with which the right answers can be detected; (2) the influence of response set; and (3) the failure to design them for purposes of occupational prediction.

PERSONALITY TESTS CAN BE BEAT

This point has been repeatedly emphasized in discussions of personality testing. It is easy to see that the right answers to questions like these can be readily detected.

Do you feel tired most of the time	Yes	No	?
Are your feelings easily hurt?	Yes	No	?
Do you make friends easily?	Yes	No	?
Are you often nervous and excited	Yes	No	?

Efforts have been made to overcome this problem in the use of the traditional type personality scale. Use of many items, only some of which are scored, is one method. Counting the use of the ? response in items using yes, ?, no is another. Excessive use of this response is taken as indicating evasion. Insertion of items to which practically everyone answers yes or no is another method. Divergence from the expected response on several of these items renders the honesty of response suspect. These methods are of the indirect and repair variety. What is needed is a form of item which is less susceptible to evasion and to outright falsification than the traditional ones.

THE INFLUENCE OF RESPONSE SET

The usual personality item requires a response on a scale of *yes-no, agree-disagree,* or *like-dislike*. A neutral or indifference point is usually included, such as: a ? in the first instance, *undecided* in the second and *indifferent to* in the third. The scale is frequently extended to five points and occasionally to seven. The different ways these can be interpreted influence the validity of the tests. Some agree if they accept the statement at all; others reserve the use of *agree* for propositions with which they are in whole-hearted agreement. Further there are consistent individual differences in tendencies to acquiesce, to use the neutral point, or to use the extremes in a 5 or 7 point scale. The evidence for these tendencies has been interestingly summarized by Cronbach (1946). Such tendencies can affect validity, raising or lowering it, depending on the relationship of these tendencies to the test score and criterion. Available findings point to the desirability of developing types of items less subject to these response sets, which, it should be added, merit study in their own right.

TESTS NOT DESIGNED TO PREDICT OCCUPATIONAL SUCCESS

One of the most efficient ways of building a test is to (1) collect a series of items which have some likelihood of being related to the criterion, (2) select the items on the basis of a high relationship to the criterion and low interrelationships, and (3) test for shrinkage in a new population. The construction of personality tests usually does not conform to principles 1 and 2, and 3 is not utilized to the extent it should be.

In the assembling of items for a personality test, attention has been focused, not on a specific prediction problem, but on adjustment concepts, particularly on presence or absence of psychological and/or psychiatric symptoms. Job analysis which has proved so useful a tool in developing aptitude and ability tests for prediction purposes, unfortunately contributes little help in assembling a series of personality items for tryout. It does not reveal what personality traits are required for a job with the clarity it points to the aptitudes and abilities involved. The relative failure of job analysis in the personality field has probably strengthened the tendency toward the adjustment orientation, an orientation which appears to have restricted the item content of such scales.

Not only has this orientation restricted the type of item tried, but it results in a scale in which the reasonably well adjusted individual finds little which applies to him. He cannot describe himself fully even if he is willing to do so. This results in mean scores on such tests which are well below those theoretically expected. Anyone answering on a chance basis will appear maladjusted. Application of such tests to problem populations who are uncooperative or poor readers may suggest an illusory validity for this reason. A further disadvantage of the adjustment centered item is the resistance to replying it arouses in many instances.

Owing to the difficulty in obtaining a criterion of dominance, introversion, emotional stability, and the like, items have been selected for a test, not on the basis of their relationship to such an external yardstick, but on the basis of the internal criterion of total score on all the items tried out. This results in a scale of items which correlate high with each other, a condition not conducive to efficient prediction. Insofar as occupational success is dependent on the possession of specific and unrelated traits the method tends to eliminate just those items in which we are most interested.

The comments concerning the method of internal consistency are conventional. While the method has the weaknesses mentioned, the failure of traditional type personality items, selected with reference to an external criterion, to meet the test of cross-validation (application to a new population) suggests it is not the method of construction but the item-content itself, the ease of detecting the right answers, or the influence of response sets that are primarily responsible for the general failure to obtain validity in predictive problems. Two types of item based on a different orientation are more difficult to beat and eliminate the influence of response set.

THE FORCED-CHOICE TECHNIQUE

ORIGIN AND DEFINITION

As the name implies the forced-choice technique involves forcing the subject to choose between two or more alternatives. In the most general sense it has been used ever since the invention of the multiple-choice item. More specific application of the principle is described by Shipley, Gray, and Newbert (1946) in the construction of an inventory designed for screening psychiatrically undesirables for the Navy. This inventory involves traditional content. The items are paired objectively on differential validity and subjectively on social desirability. The technique to be described here involves a radically different item content as well as a method for objective pairing of alternatives to make the right answers difficult to detect. The method was invented by R. J. Wherry and is reported here as used by the Personnel Research Section, AGO, in 1945 and later.

WHERRY'S TECHNIQUE

Wherry started with the notion that the way to prevent detection of the right answers to personality items was to present pairs of self-descriptions, both members appearing of equal value to the subject, but having differing significance for the purpose at hand. The problem was to develop a method by which these ends could be achieved. The first step was to collect a series of descriptive phrases which could be scaled. Lists of words descriptive of personality were first consulted. Factor studies were reviewed to insure as complete a coverage as possible with a minimum of duplication. Descriptive words and phrases like *leader, hold to a purpose, industrious, thrifty,* and *capable* were used.

The second step was to scale the single descriptions to discover which could be paired to appear of equal value to the subject. This step will be illustrated as it was accomplished by the Personnel Research Section in the construction of the Army Biographical Information Blank (BIB) for use in selecting the better officers for retention in the peacetime Army (AGO, Report 703). Wherry's list of 570 words and phrases were presented to a group of 200 officers (subjects used for this scaling should, of course, be like those to whom the final test is to be administered) with instructions to indicate how well they would *like to be described* by each. The scale used was:

Like extremely
Like strongly
Like mildly
Indifferent to
Dislike mildly
Dislike strongly
Dislike extremely

Values from 0 to 6 were assigned to these phrases and the mean and a measure of dispersion for each description were computed. Items with large dispersions were not used. This mean is the preference value or index. Two descriptions with the same preference index presumably are equally acceptable to the subject.

The third step was to scale the descriptions in terms of their significance for the purpose involved. The same list for which preference values were obtained were given to an additional group of officers with instructions to indicate *how desirable each trait was for success as an Army officer*. The scale used:

Extremely desirable
Strongly desirable
Mildly desirable
No effect on success
Mildly undesirable
Strongly undesirable
Extremely undesirable

Again values ranging from 0 through 6 were assigned and means computed. These are the discriminative indices.

The fourth step is to pair the descriptions so that the members of each pair appear equally acceptable but differ in significance. This is done by searching for descriptions which have the same preference index but which differ in discriminative index. Examples of possible pairings on the basis of AGO data are given below:

Description	*Preference Index*	*Discriminative Index*
	Pair I	
Attractive	4.80	4.50
Demands perfection	4.77	5.04
	Pair II	
Avoids advice	1.69	1.51
Puts on airs	1.63	2.02

This method of pairing should result in alternatives equally acceptable to the subject but which differ in their significance for prediction.

From this point on, the method corresponds with the steps of any good test construction program—tryout of the pairs on a criterion group, item analysis, selection of the items, and crossvalidation of the final scale.

PROBLEMS IN TECHNIQUE

There are a number of problems concerned with the use of this technique which have been only partly investigated.

1. As an alternative to starting with a list of descriptives culled from the literature, it is possible to obtain anonymous essays concerning persons known to the essayist (but not to the experimenter) to be high or low on the criterion concerned. Appropriate descriptions can be selected from these. The main advantage of the method is that it insures using the language of those taking the test.

2. The descriptions may be presented in triplets, tetrads, or quintets as well as by pairs, as in the originally devised method. An objection to the paired method of presentation is that it arouses resistance on the part of the subject when socially unacceptable descriptions are used such as brutal and ugly, to use an extreme example. In the construction of the BIB two methods were tried for overcoming this resistance. One method did not work. It was a quintet presentation using socially undesirable descriptions but having the subject respond in terms of someone else. This mild use of the projective principle aroused no resistance but unfortunately it retained no validity. The method that worked was presenting the alternatives in quintets such as:

A) attractive
B) avoids advice

C) avoids fatigue
D) demands perfection
E) puts on airs

[*Note:* In presenting a group we have arranged the alternatives alphabetically.]

The subject makes two responses, first, indicating which alternative is most descriptive of himself and, second, which one is least descriptive. The quintet method of presentation was based on the belief that the unpleasant tone of least desirable descriptives would be reduced by (1) including them with the most desirable, (2) having the subject respond on a "least descriptive" basis, and (3) by the inclusion of a neutral term. The method was highly successful (AGO, Report 701) and permitted the use of such unfavorable terms as these: rough in manner, awkward, sluggish, hard, gabby, crafty, rash, and artificial in manner. The quintet method has the advantage of utilizing more potentially discriminating terms.

How to determine which five descriptions to group together is the crux of the problem. Basically it is a matter of finding two pairs and a relatively neutral term. The method of obtaining pairs has already been described. The method used in grouping two pairs in the construction of the BIB was to order the pairs according to the preference values and divide them into two halves. The first pair in the upper half is placed with the first pair in the lower half; the two second pairs are grouped, and so on. This procedure insures a rough degree of similarity in the difference between the preference values of the pairs in the several quintets. Four highly desirable words or four highly undesirable words are prevented from occurring in the same quintet by this procedure.

One could stop at this point and present the groups as tetrads. This procedure has been used in the construction of a forced-choice rating scale and worked adequately. It is my belief, however, that in the test as opposed to the rating situation, the addition of a fifth term is an advantage. It increases the probability of finding valid responses (many neutral items become scored items in the final scale), and it aids in taking the sting out of such unpleasant descriptions as are used.

This fifth term is added by finding a description with a preference value midway between those of the two pairs and with a discriminative value likewise between the discriminative value of the two pairs. A possible quintet is given below to illustrate the method.

Description	*Preference Index*	*Discriminative Index*
Attractive*	4.80	4.50
Avoids advice**	1.69	1.59
Avoids fatigue***	3.24	3.15
Demands perfection*	4.77	5.04
Puts on airs**	1.63	2.02

* Relatively desirable pair
** Relatively undesirable pair
*** Neutral item

In grouping the alternatives the main non-statistical consideration to keep in mind is the avoidance of opposites. Inclusion of opposites in the same quintet will tend, of course, to bring the group back to the very pattern the item was designed to avoid, namely the *yes-no* pattern which is so easy to beat.

3. Obtaining the response by ranking the alternatives instead of requiring indication of most descriptive and least descriptive may be a more advantageous method. The latter was used in the BIB to save space on the answer sheet and get more item coverage. In hand-scored tests the ranking method will yield all that the most descriptive technique will, and perhaps more. Ranking is therefore recommended until it is determined that no advantage is gained. The ranking method has been tried with Owens-Illinois factory girls and found to be readily understood.

4. An alternative method of scaling the descriptions to permit proper pairing is to score them directly against the criterion to obtain discriminative values and to determine preference values from the frequency with which the group actually uses the single terms in describing themselves. Use of a known criterion makes it impossible to retain anonymity. Preference and discriminative values may be distorted. Whether the additional labor of obtaining a criterion group at this stage of the method will pay sufficient dividends is not known.

ADVANTAGES OF THE FORCED-CHOICE TECHNIQUE

1. The basic advantage to the forced-choice technique is that it is difficult to detect the right answers to items constructed by the method described.

2. The use in the scaling procedure of a wide variety of personality descriptions gets away from limitations imposed in approaching the problem from the standpoint of adjustment or clinical symptoms that have been the content of traditional personality tests. What traits are of importance will emerge through a study of the scaling by those in the best position to know. Until we know more of the impor-

tant dimensions of personality this is a considerable asset.

3. Data for pairing the items can be obtained without the use of criterion groups. The data can be obtained anonymously. All that is necessary is that the preference values be obtained from a group like that to which the final test is to be given and the discriminative values be obtained from those familiar with the personality requirements of the job.

4. The form of presentation and response required reduces contamination by response sets. Tendencies to acquiesce, to take neutral or extreme positions on a scale lacking well defined reference points, cannot operate. It eliminates whatever set is basic to the finding that negatively stated items tend to be more valid as well as yield means nearer the favorable end of the scale (Rundquist, 1940). No evidence was uncovered to indicate that the less favorable descriptions were more or less valid than the favorable ones. Likewise no tendency for the most and least descriptive responses to behave differently was discovered.

The technique does not, of course, eliminate a set on the part of the subject to answer the items on some other basis than describing himself—a difficulty found in all forms of personality measurement, including projective techniques and the interview. It does, however, increase the difficulty of finding another basis on which to reply.

5. The technique forces a real discrimination by the subject. One can easily convince himself of this by a little introspection while responding to traditional personality questions and to those of the kind described. It may well be that the forcing of a real discrimination is the heart of the technique.

6. The kind of item used—painstaking, overactive, serious-minded, sincere, selfish—is less objectionable and more applicable to the normal person taking a personality scale than is the kind of item used in general adjustment or neurotic inventories.

7. The method of item construction increases the efficiency of finding discriminative items. Evidence for this will be presented later.

BIOGRAPHICAL INFORMATION AS A PREDICTOR

The importance of personal background—community, home, education, health, occupation—has received so much emphasis as a determiner of personality characteristics that it is surprising that interest in its use for specific prediction problems has been so slow in growing. The social case history represents the most extensive collection of such information. When collected in this form its primary use is in understanding and dealing with problems of personal adjustment. The interpretation is based on general psychological and psychiatric concepts. In much more limited fashion, application blanks and interviews have collected this kind of information and interpreted it in an even more subjective manner.

Several efforts have been made to use personal background information as a predictor of occupational success, particularly for salesmen (Goldsmith, 1922; Kornhauser and McMurry , 1941; Kurtz, 1941). All report some success. It has been found to predict college entrance (Bittner, 1945). H. M. Johnson (1946) in a recent review of the literature on the detection and treatment of accident-prone drivers points out that biographical information is one of the unexploited and most promising predictors in this area of behavior. More thorough exploitation of this type of item has occurred during World War II, the Army, the Navy, and the Air Forces having constructed biographical inventories. The fullest exploitation requires coverage of all the areas of experience covered in a good case history.

The major problem in constructing the items is objectifying the method of obtaining the information and reducing it to scorable form. The items must have alternatives which permit a response from persons of widely differing experiences. In the Army BIB, for example, provision had to be made for those with and without college experience, and with and without civilian job experience. Grouping items pertaining to a particular type of experience is one solution. Addition of an alternative, such as, "I did not go to college," to all items concerned with educational background is another method. In addition to making provision for all possible responses great care needs to be taken with the stem of the item. For example, if the item concerns the course taken in high school, provision has to be made for those who changed; if it concerns parents, what to do if brought up by someone other than own parents must be indicated. In brief, biographical items must conform to the principles of good item writing, except that it is more difficult to make them do so than it is, say, for intelligence and achievement test items.

Biographical items can cover a great deal of material in a short space. They are not limited to strictly factual material, but can cover preference, attitudes, and interpretation of experience. The items are acceptable to the subjects because they appear reasonable to them. Their greatest advantage is in relating experience directly to occupational success rather than indirectly through the interpretation of an interviewer. It becomes possible to identify elements of experience that are related to specific jobs. Use of a number of such items, weighing each in terms of its contribution to the prediction of success on a specific job, is clearly preferable to the subjective weight given by an interviewer who can hardly avoid weighing certain items heavier than others or making a stereotyped application of the same bit of information to all jobs. This point of view is an expression of my belief that the same items will be scored differently for different jobs, evidence in support of which will be reported later. It can hardly be otherwise in the industrial situation where applicants are self-selected in ways to alter the significance of particular types of information and where personality requirements probably vary widely from job to job and may even differ for the same job in different settings.

Like items constructed by the forced-choice technique, biographical items are not affected by response sets, nor are the right answers easy to detect.

Various techniques of presenting biographical items are presented below:

Did you make a first or second string high school team in any of the following sports? (Mark as many answers as apply.)

A) no
B) yes, in football
C) yes, in basketball
D) yes, in baseball
E) yes, in track and field events
F) yes, in boxing
G) yes, in tennis
H) yes, in swimming
I) yes, in wrestling

How old were you when you learned to use a shotgun or any gun of .22 caliber or larger? (Mark only one answer.)

A) younger than 10
B) 10–13
C) 14–17
D) 18–20
E) 21 or older

What was your parent's attitude about your choosing the occupation you wanted to follow? (Mark only one answer.)

A) My parents encouraged my occupational interests, but the choice was up to me.
B) My parents attempted to force me to accept the choice they had made.
C) My parents attempted to influence me to accept their choice, but didn't interfere with plans I made.
D) My parents didn't care what I decided to do.
E) My parents and I have agreed on what I would do when I was very young.

How large was the student body of the college or university you attended? If you attended more than one, describe the one you attended longest. Do not count colleges attended as AST or ASTR trainees. (Mark only one answer.)

A) less than 500 students
B) between 500 and 1,000 students
C) between 1,000 and 4,000 students
D) between 4,000 and 8,000 students
E) More than 8,000 students

What was the highest cadet rank you held in your college ROTC unit? (Mark only one answer).

A) private
B) noncommissioned officer
C) commissioned officer
D) My college offered optional ROTC training but I did not take it.
E) My college did not have an ROTC unit.
F) I did not go to college.

To which of the following activities have you devoted considerable time and energy because of your interest in them? (Mark as many answers as apply.)

A) amusements (dancing, movies)
B) art
C) music
D) collecting (stamps, coins)
E) dramatics
F) literature
G) journalism
H) mechanical and manual activities
I) Boy Scouts
J) photography
K) radio
L) scientific activities
M) ROTC
N) semireligious activities
O) none of the above

GENERAL PROBLEMS

A brief discussion of the Army Biographical Information Blank (BIB) will serve to illustrate certain general principles of personality test construction and the value of the two techniques of item construction that have been described. The BIB differs radically in philosophy and method of construction from the traditional personality scale.

1. The approach was to be empirical. No theories concerning traits required by a good officer were to limit the item content. As many items of as many kinds as would be practical to try out would be included.

2. Two types of items were excluded:
   a. The traditional item with yes-no response. It was not felt that these had any promise in predicting officer competence.
   b. Items of extreme projective nature. This course was taken partly because of time limitations and partly because of the absence of any convincing evidence concerning the usefulness of this form of item in prediction studies.

3. The forced-choice principle would be exploited. The fact that a great deal would depend on the results of the BIB made it essential that the possibility of beating the test be reduced to a minimum.

4. Cross-validation would be emphasized both in item selection and score validation.

A wide variety of items were tried out, more than 1,100 in all, with a wide variety of techniques. Space limitation prohibits describing the various techniques tried. The major findings concerning type of items are these:

1. The biographical items (which cover Army and civilian job experiences, educational and health history, leisure activity, preferences and interests, home and family background, social and community relations) and the self-descriptive quintets constructed by the forced-choice technique turned out to be the most valuable, and constitute the major portion, of the 204 items in the final form of the scale.

2. Biographical and forced-choice self-descriptive quintets were most efficient: 65.3 per cent of the former tried out survived intensive validation; 69.3 per cent of the self-descriptive quintets.

3. The efficiency of the forced-choice items varied with the content and/or the basis for pairing. Of item pairs with traditional content, taken from the Minnesota Multiphasic and paired (AGO, Report 700) on the basis of an adjustment criterion, 21.8 per cent survived validation. Of a series of miscellaneous pairs for which the only basis of pairing was guesswork, some traditional in content and some not, 13.3 per cent survived. Even allowing for possible influence of difference in number of alternatives, these results contrast unfavorably with those presented under 2 above.

4. Items presented on a scale subject to the influence of response sets failed. Of a series of attitude items and a series of annoyance items, both adapted to the Army situation, the first presented with the *strongly agree, strongly disagree* scale and the second with a scale running from *extremely annoyed* to *not annoyed,* so few survived validation as not to warrant space for special directions.

RELIABILITY AND VALIDITY

The details of the validation cannot be described and must await publication by the Personnel Research Section. It must suffice to point out that the criterion was much better than the usual one and that the amount of cross-validation, both in item selection and scale validation, exceeds anything heretofore published. The validity (AGO, Report 704) of the scale as a whole was found to be 0.32 (N=689) and 0.38 (N = 830). A few items were eliminated from the form on which these validities were obtained and the final form which was used in the officers retention program was checked for validity on a group of 1,344 officers. The correlation was found to be 0.33 (AGO, Report 716). In this sample, the percentage of good, poor, and average officers included in certain score ranges is:

Score	Percentage of Total Group	*Criterion Group*		
		Low	Middle	High
100 and above	32.5%	22.8%	22.4%	54.7%
90-99	32.1	32.4	28.5	39.1
89 and below	35.1	49.3	31.2	19.5

Considering that predictions of intangibles like officer competence have previously been outside the scope of objective personality measures, these results are highly promising. The techniques that made them promising deserve extensive study and application.

The best test of reliability is whether a test does consistently what it purports to do, that is, whether it selects the same proportion of good, poor, and average persons for specified score ranges for different populations. This test the BIB meets well. The proportions in other samples are very close to those presented above. This type of reliability is considered more important than any coefficient dependent on internal consistency. Dividing the BIB into equivalent halves on the basis of judgment yielded a reliability of 0.78 (AGO, Report 704) as estimated from the Spearman-Brown formula. This is undoubtedly an underestimate. Reliability and validity are highly relative problems as Guilford (1946) has recently emphasized. What is acceptable for either depends on the particular purpose involved.

While the BIB has been most thoroughly validated for officers, it has been tried in a limited way in officer candidate schools. In one school for two successive classes (AGO, Report 711) it yielded a validity of 0.45 and 0.55 (*N*'s-40–50).

The validity of the separate sections of the BIB has not been computed separately for populations uncontaminated by item selection procedures. Wherry has informed the writer that he has obtained validation in the neighborhood of 0.50 in predicting success of foremen in industry for self-description pairs.

SPECIFICITY OF PREDICTION

The first sifting of the items of the BIB was done on the basis of officer candidates; the final selection on the basis of officer responses. It became clear that the scoring key developed on the basis of officer candidates could not be used for officers if maximum validity were to be achieved. The validation reported above for officers and officer candidates are for the same items, but with overlapping but different scoring keys. The extensive validation procedures in item selection precludes chance as the explanation. Whether the specificity is a result of differences in the criterion evaluation of personal leadership qualities by younger (inexperienced) and older (experienced) men, or to differences in the way the words, particularly of the self-descriptive items, are evaluated is not known. Selection imposed by elimination of unsuccessful officer candidates may also operate. To use a familiar example: if either self- or imposed-selection brings applicants, all of whom are sufficiently intelligent to do the job, the use of an intelligence test will not aid in further selection. In just such a way, items of personal background can alter in significance through subtle factors in selection.

DISCUSSION AND SUMMARY

While the importance of personality factors in occupational success has been uniformly emphasized, attempts to measure these factors have by and large failed. The ease in detecting right answers to personality test items, the considerable role of response sets in determining the way in which they are answered, and their concern with adjustment factors rather than specific prediction problems are three major reasons for this failure. Of these the first is paramount. At least it obscures the evaluation of the others.

In this section the emphasis has been on describing two approaches to personality measurement which hold real promise for the future. The one, the forced-choice technique, was designed specifically to prevent detection of the right answers to personality items. It is immune to contamination by response sets inherent in the item of the traditional type. The technique brings out the aspects of personality as revealed by self-description that are significant for a given job and reduces reliance on theoretical systems of personality organization. Many questions concerning the technique and the best way of applying it are in need of research. But a small beginning has been made in the exploitation of its potentialities.

The other approach, the use of biographical information, by its nature is little affected by response sets, nor is it subject to easy detection of the right answers.

Used in the same instrument these types of items obtained moderate and consistent predictions in a leadership area—officer competence—an area in which prevailing opinion is that prediction is impossible except possibly by clinical judgment taken on faith as valid.

The methods described are empirical. The forced-choice technique, in particular, used with self-descriptive items that are applicable to the reasonably well adjusted provides information on what to include in a test. Until, through such methods as factor analysis (Cattell, 1945), there are available dimensions of personality to guide in personality test construction, the failure of traditional type personality items in prediction studies, leave us no choice but to resort to empirical methods of the kind described. It is one of the merits of the forced-choice technique, as here described, that it indicates what should be included.

These techniques have much to contribute to occupational selection. How much they will contribute to guidance depends on further findings concerning specificity of prediction. If it should turn out that prediction is specific to particular companies, even for similar jobs, the difficulty of use in guidance is evident. When the dimensions of personality become better known so that subtle factors in selection can be better evaluated, possibilities of use for guidance will be considerably brighter.

The results obtained through the use of these techniques offer the promise of extending our understanding and prediction in areas, such as leadership, which are usually considered outside the range of objective techniques. Reliance on the interviewer as the major method of estimating personality characteristics can be reduced. The view that the interview should be limited to securing information not obtainable by objective means is quite different from that implied in the usual statement, "Tests are a means of gaining facts to aid the interviewer in arriving at a final judgment." Both tests and the interview are predictive tools and should be evaluated in the same terms. In one situation the interview, with or without test information, may be the best predictor. In another it may contribute nothing and tests carry the burden of prediction. To treat all tests as aids to the interview in the selection process runs the great risk of having known validities destroyed. The possibilities of the effective uses of the interview in prediction are just beginning to be explored. There is little point in assigning it tasks that can be done better and less expensively by other means.

Our concern in this chapter has been with methods of personality testing having promise for predicting occupational success among the reasonably well adjusted range of people. This problem, while clearly not unrelated to those of the clinic where the major concern is the understanding and alleviating maladjustment, is equally clearly not the same. Considering the range and variety of human experience and the considerable amount of specificity in it, it becomes understandable why tools developed for the latter purpose, emphasizing relatively few basic drives and reaction tendencies, do not meet the needs of the former situation. The reverse likewise holds. As pointed out earlier, concern with problems of adjustment, important as it is, has tended to restrict thinking about the type of item to be used and hence has been a delaying influence in the development of techniques that show promise in occupational selection and guidance.

READING REFERENCES

Achilles, P. A., "The Consulting Services of the Psychological Corporation." *Journal of Consulting Psychology,* 1946, **10,** 120–126.

Adjutant General's Office, Personnel Research Section, War Department, "Construction and Selection of Items for the Biographical Information Blank (BIB)," *Report* No. 703, 1945.

———, "Item Analysis of the Multiphasic Personality Inventory," *Report* No. 700, 1945.

———, "Methodological Investigation of the Forced-Choice Technique," *Report* No. 701, 1945.

———, "Prediction of Leadership Qualifications of Officer Candidates in the Signal Corps," *Report* No. 711, 1946.

———, "Validation of Form E of the Biographical Information Blank (BIB)," *Report* No. 716, 1946.

———, "Validation of a Program for the Selection of Officers for Retention in the Peacetime Army," *Report* No. 704, 1946.

Bittner, R. H., "Quantitative Predictions from Qualitative Data: Predicting College Entrance from Biographical Information." *Journal of Psychology,* 1945, **19,** 97–108.

Cattell, R. B., "The Description of Personality Principles and Findings in a Factor Analysis." *American Journal of Psychology,* 1945, **58,** 69–90.

Cronbach, L. J., "Response Sets and Test Validity." *Educational and Psychological Measurement,* 1946, **6,** 475–495.

Goldsmith, D. B., "The Use of the Personal History Blank as a Salesmanship Test." *Journal of Applied Psychology,* 1922, **6,** 149–154.

Guilford, J. P., "New Standards in Test Evaluation." *Educational and Psychological Measurement,* 1946, **6,** 427–438.

Hunt, W. A., and I. Stevenson, "Psychological Testing in Military Clinical Psychology: II. Personality Testing." *Psychological Review,* 1946, **53,** 107–115.

Johnson, H. M., "The Detection and Treatment of Accident Prone Drivers." *Psychological Bulletin,* 1946, **43,** 489–532.

Kornhauser, A. W., and R. N. McMurry, "A Scientific Procedure for the Selection of Salesmen." *Journal of Applied Psychology,* 1941, **24,** 11–17.

Kurtz, A. K., "Recent Research in the Selection of Life Insurance Salesmen." *Journal of Applied Psychology,* 1941, **25,** 11–17.

McMurry, R. N., "Validating the Patterned Interview." *Personnel,* 1947, **23,** 263–272.

Rundquist, E. A., "Form of Statement in Personality Measurement." *Journal of Educational Psychology,* 1940, **31,** 135–147.

Shipley, W. C.; F. E. Gray; and N. Newbert, "The Personal Inventory—Its Derivation and Validation." *Journal of Clinical Psychology,* 1946, **2,** 318–322.

30. EFFECTIVENESS OF SELECTION DEVICES

MARION W. RICHARDSON, PH.D.

Director of Research and Chairman of the Board, Richardson, Bellows, Henry and Company, Inc.; Diplomate in Industrial Psychology, American Board of Examiners in Applied Psychology; Member, Editorial Staff, *Psychometrika; Educational and Psychological Measurement* and *Personnel Psychology;* President: Personnel Psychology, Inc. Formerly: Chief, Personnel Research Section, AGO, Department of the Army; Chief, Test Development, and Principal Examiner, US Civil Service Commission; Examiner, Board of Examinations, University of Chicago; Supervisor of Selection Research, The Procter and Gamble Company; President, Psychometric Society.

THE PROBLEM

How good is a test, a battery of tests, an interview, an application blank, or any other device used for selection of personnel? One answer is in terms of the "judgment" of those concerned with the use of such devices. Such judgments, whether made by technically competent personnel or by practical operating people are more apt than not to be based on insufficient or nonpertinent information. Furthermore, they are usually influenced strongly by the special enthusiasms and prejudices of the judge. Thus some "practical" personnel men have accepted estimates indicating that tests have much less than perfect predictive value for selecting personnel, without insisting that the interviews and procedures they have used for selection be subjected to the same criterion. According to this point of view anything new or unusual must be subjected to strict trial, but the time-honored procedures need not be subjected to such severe scrutiny. Some professional workers, obsessed with the clinical point of view, think in terms of individual cases, and see no particular need of any type of statistical check-up. Yet, an accumulation of "cases," clinical or otherwise, constitutes a sample population that can be checked on, and must eventually be checked on, if any real progress is to be made. The comparative ease of statistical check-ups of selection procedures has perhaps caused certain of us to neglect such considerations as these: (1) The performance of a trainee in a training course is a predictive element, and can be treated in the same manner as an aptitude test; and (2) a preliminary requirement for an applicant, such as high school graduation, is a predictive element, the effectiveness of which can be checked by statistical procedures.

The problem that comes to the front is just how shall the effectiveness of any one, or a combination of several, of these procedures be evaluated. The results must be intelligible to the person who does not think in terms of the academic statistical concepts, but who nevertheless wants a quantitative answer. The present section is concerned with this problem. It is, of course, specifically recognized that certain methods of evaluating the effectiveness of selection devices have neglected such matters as the cost of training personnel, the man power available for the position under current labor markets, and the value to the company of efficient employees over a period of several years.

CONVENTIONAL TECHNIQUES

The predictive value of a selection device has an immediate but highly incomplete and somewhat misleading answer in the standard error of prediction. When the predicting device (which may be a single test, interview, procedure, or any linear combination thereof) is correlated with a criterion of success on the job, a coefficient of correlation (simple or multiple) is obtained. This correlation indicates the degree of concomitant variance shared by predictor and criterion. The variance in common is proportional to r^2 (r = the correlation coefficient) and the unpredicted variance is proportional to $1 - r^2$. As customarily expressed, the standard error of estimate is given by

$$\sigma_E = \sqrt{1 - r^2} \cdot \sigma_c,$$

where σ_c is the criterion variance. By this reasoning, a validity coefficient of $r = 0.71$ indicates that one half of the total variance is accounted for by the predictive battery. This measure means that a perfect measure would give a perfect rank-order of all personnel tested or otherwise assessed, so that the person having the highest rating on the test battery would eventually be the best worker, the second highest, the second best, etc., to the lowest man in the distribution. This measure is unduly rigorous, since it includes variance within acceptable personnel, and what is more important, variance within personnel quite unacceptable in any circumstances. In other words, if the selection battery predicts that a man will fail on the job, it is not required that it describe just how bad a failure he will be. This procedure may underestimate the value of a selection predictor.

DICHOTOMOUS GROUPINGS

One effort to overcome the difficulties just described has been to express successive intervals or subranges of the prediction in terms of the probability of success, where the criterion is susceptible to division into "successful" and "unsuccessful." This type of analysis implies that such dichotomy of the criterion has meaning and relative stability. A typical expectancy table might be:

Predictor Scores	*Probability of Succeeding, in Percentages*
10 – 19	2.4%
20 – 29	4.3
30 – 39	16.4
40 – 49	38.2
50 – 59	63.4
60 – 69	75.8
70 – 79	83.6
80 – 89	86.2

By use of such a table, reasonable test predictor standards can be set high enough to effect a real improvement in quality of personnel, without sacrifice of needed numbers. Furthermore, from such data it is possible to calculate directly the increased effectiveness attained by the application of any given standards.

The expectancy table technique is, however, only a variant of the correlation technique. Indeed, it can be shown that the results are readily convertible to correlation coefficients under certain limiting circumstances (Richardson, 1936). The expectancy table has the advantage that the critical score can be set realistically in terms of the labor market, and that it carries in itself the generalization that the validity is always less than perfect.

A RATIONAL TECHNIQUE FOR COST ESTIMATION

Another approach to the problem of evaluating selection procedures involves placing a value in terms of money saved or other measures of increased efficiency.

FORMULA

A convenient measure of the value of a selection test or other device is given by the formula:

$$E = \frac{r(1 - P)(k - 1)}{P(k - 1) + 1}$$

E = percentage increase in effectiveness due to the selection procedure. Thus if E turns out to be 8 per cent, it means that for every hundred employees selected, the effectiveness of 108 is realized, or 8 per cent of productive work is gained without increase in the payroll. The value of E is calculated by substituting actual numbers in the formula. r = the validity coefficient of the test, test battery or interview as determined by the ordinary statistical procedures. P = the percentage of applicants selected. It must be set in relation to the labor market. A convenient percentage to adopt as a first approximation is the percentage of present employees who are regarded as very satisfactory or better. P must not be less than 10 per cent or more than 90 per cent. The factor k is the ratio of the average effectiveness of "satisfactory" to "unsatisfactory" (or less satisfactory) employees. This can be obtained by dividing the average production of very satisfactory employees by the average production of unsatisfactory employees, with no middle class. If no production records are available, consensus of judgments of supervisors may be utilized, or other estimates made. The exact value of k varies with the job. It may be very large in high grade jobs; in most jobs it is not often less than three.

EXAMPLE. A test battery has a validity of $r = 0.50$. Records indicate that the upper 20 per cent of a group tested produces one and one half times as much as the lower 80 per cent.

$$E = \frac{.50(1 - .20)(1.5 - 1)}{.20(1.5 - 1) + 1} = 18\% \text{ approximately}$$

SPECIAL CHARACTERISTICS

The foregoing formula points out the essential characteristics of any selection device, viz., (1) *The percentage increase in effectiveness is directly proportional to the validity coefficient, and* (2) *any selection* *device results in greater improvement in selection when the proportion selected is small rather than large.*

The formula is easy to use, and when applied with conservative values of k, will give the nonstatistical person a realistic description of what can be expected by standardized selection procedures, including tests, interviews, and personal history.

READING REFERENCES

Peters, C. C., and W. R. van Voohis, *Statistical Procedures and Their Mathematical Bases.* New York: McGraw-Hill Book Company, Inc., 1940.

Richardson, M. W., "The Relationship between the Difficulty and the Differential Validity of a Test." *Psychometrika,* 1936, **1**, 33–49.

CHAPTER V

FIELDS OF PERSONNEL PSYCHOLOGY
SELECTION AND TRAINING

31. Clerical Personnel	W. Leslie Barnette, Jr.
32. Mechanical Personnel	Thomas W. Harrell
33. Field Salesmen	Marion A. Bills
34. Retail Sales Personnel	Kate Lewis
35. Management's Use of Personnel Research in Training	Roger M. Bellows
36. Federal Civilian Personnel Administration	L. J. O'Rourke
37. Military Personnel Management	Donald E. Sisson
38. Metropolitan Police Personnel	Sadie Myers Shellow
39. Subnormal Workers	Theodora M. Abel and Emily Burr
40. Employment Agencies	Lorin A. Thompson

31. CLERICAL PERSONNEL

W. LESLIE BARNETTE, JR., PH.D.

Department of Psychology, New York University.

In the field of business perhaps the greatest progress in testing and job standardization has been made in the clerical occupations. The reasons are that job demands are similar; research agencies cooperate in making surveys and job analyses, as well as in constructing and validating standardized tests. This work has been particularly useful, not only because large numbers of applicants pour into the white-collar labor market, but also because occupational choices of students show that a large proportion prefer clerical positions. Too often commercial high schools, business schools, and counselors assume, contrary to fact, that all students have the necessary aptitudes to be trained for and can subsequently be placed in office positions. Personnel psychology, with its techniques and instruments for sound selection, guidance and training of workers, can do much to divert this flow elsewhere.

Anderson and Davidson (1940) in their volume on occupational trends in the United States show that, for the nation as a whole, a stabilized level in the employment of clerical labor had about been reached. This group constituted 4 per cent of the entire American working population in 1910 and 7.8 per cent in 1930. The 1940 census data showed a still further increase. It therefore appeared likely that the saturation point, particularly for stenographers and typists, was being neared. The war years, however, radically changed this situation. Attendant with the great expansion of government activities and the increased bookkeeping loads placed upon most business firms, well trained stenographers and typists became a rarity. The shortage continues to the present time. Increased mobility of clerical personnel has thus received more attention and has been increasingly demanded by employers; this mobility, in turn, is largely a function of more comprehensive school training oriented specifically to current business needs. What is needed, as Bell (1940) has pointed out, is a general clerical and business curriculum for clerical workers which will lay the foundation for increased mobility rather than the merely specialized instruction in typewriting and stenography still so prevelant in most public schools.

The responsibility of clerical employees is frequently greater than is generally assumed. They are often called upon to make decisions of a minor executive nature. Planning is not exclusively an executive function: 30 per cent of the clerical force of the Aetna Life Insurance Company carry out duties of a supervisory nature (Bingham, 1935). Pond (1941) emphasizes that the effect of technological improvements in office work has been to reduce the number of repetitive tasks, with the result that clerical workers must be more capable of analytic thought than heretofore has been the case. It is thus an error to think of the clerical occupations as principally routine.

SELECTION BY TESTS

Many companies have found psychological tests valuable in selecting clerical workers. Moore in a National Industrial Conference Board report (1941) conservatively estimated that between 300 and 400 companies hiring 500 or more workers were using psychological tests for worker selection, clerical and otherwise. The Dennison Manufacturing Company, in one series of tests, found that 90 per cent of those making high scores were satisfactory employees; among those making low scores, less than half became good workers (Dennison, 1931). The Aetna Life Insurance Company, with tests used since 1925, found a consistent relation between scores at the time of employment and the grade of job attained later by the worker, i.e., scores were higher as the job level increased. Bartlett (1938) described similar findings with a battery of clerical tests used by the Atlantic Refining Company. The above-mentioned National Industrial Conference Board report outlines the experience of other companies (Scovill Manufacturing Company, R. H. Macy, Detroit Edison, and others) with psychological tests.

Besides serving as a check on the experience and training background of the applicant, psychological tests disclose latent abilities. Wadsworth of the Pacific Lighting Company (1939) reported that about a third of their line supervisors first attracted the attention of the personnel department through the use of psychological tests.

Tests, then, save time, contribute to better promotion procedures, and aid in the guidance of employees. The guidance aspect needs more emphasis. Stott (1935), reporting from Great Britain, found no significant relation between satisfaction and salary, age, experience, and leisure interests in secretarial work; 38 per cent chose their jobs because it was "the only thing left open to them." Deemer (1944) in a follow-up study of approximately two hundred Massachusetts high school students trained in stenography found that 60 per cent who reported employment were unable to secure jobs where they could use shorthand.

Surveys disclose that there is need for wider usage of tests in clerical employment. In Bell's (1940) Baltimore report, one finds a ratio of best to worst clerks (121 card-punch operators) of 1:10; for eighty calculating machine operators, 1:3.8. Yet the 1936 report of the National Industrial Conference Board (the most recent in which statistical data are presented) showed that only 5.8 per cent of 2,452 companies (involving more than 4,500,000 workers) used tests to select their clerical employees. Insurance companies have been decidedly in the lead here. Larger companies, probably because they have more standardized tasks, were found to use tests more frequently than the smaller ones.

Workers are generally suspicious of tests. Their introduction into a company's personnel practice has to be accomplished tactfully, especially if the tests are to be administered to employees already on the payroll. It is well to discuss the problem with the employees at the outset, emphasizing that their abilities are already known, that test scores can have no bad effect on job tenure and that the tests may disclose latent abilities.

JOB ANALYSIS

A job analysis must be made before instituting any testing program. In the clerical field such analyses are usually of the descriptive or enumerative type in contradistinction to the time and motion variety. Nichols (1927) reports an analysis of 34,513 office employees in fifty-four firms in seven different industries. Eleven essential traits were defined. The top four were: systematic habits, accuracy, concentration, manual dexterity. Cooperation and mental alertness were rated somewhat lower. Last in this grouping came courtesy, accuracy in details and figures, willingness to assume responsibility, judgment, clear thinking. A more recent survey would undoubtedly disclose a different rank order in view of the increasing complexity of many clerical jobs today. Charters and Whitly (1924) listed and ranked 871 secretarial duties, among which the most important were: accuracy, responsibleness, dependability, intelligence, courtesy, initiative, and judgment. Bingham (1935) defined four kinds of clerical aptitude which he considered valuable: (1) perceptual, (2) intellectual, (3) motor and (4) skills susceptible to training (as spelling, multiplication, etc.).

The Life Office Management Association has made extensive analyses of clerical jobs as found throughout the country in life insurance companies. The assumption here was that all clerical work could be reduced to three elements: doing, checking, supervising. Nine classes of jobs were analyzed and 107 different operations defined as of 1940. This type of job evaluation is limited to large companies and is most easily applied to low-level positions; one also is apt to encounter union disagreements and different company personnel practices. It would seem best not to attempt industry-wide standardization when jobs themselves and management policies are as divergent as they are currently found to be (Moore, 1944).

OCCUPATIONAL ABILITY PATTERNS

The development of occupational ability patterns is a recent approach to job analysis. Such patterns are derived from the mean scores made on various standardized tests by successful workers engaged in a particular occupation. Figure 1 presents a comparison of the patterns of clerical employees and retail saleswomen in both standard and percentile scores. Women office clerks here were stenographers, typists, file clerks, and general office workers. Figure 1 shows that this group does best on tests of abstract functions (educational and clerical aptitude). Dvorak (1935) has reported on the relation of occupational ability patterns to clerical success. Women office clerks were separated into five groups according to rated efficiency. Test scores of the least satisfactory groups were found to be proportionately lower than those of the more efficient. Such patterns are approximately the same regardless of place of employment. Typicalness, rather than trait superiority, is important for success in an occupation.

CLERICAL TESTS

Clerical tests do not measure all the traits demanded by job specifications, but most such tests do locate

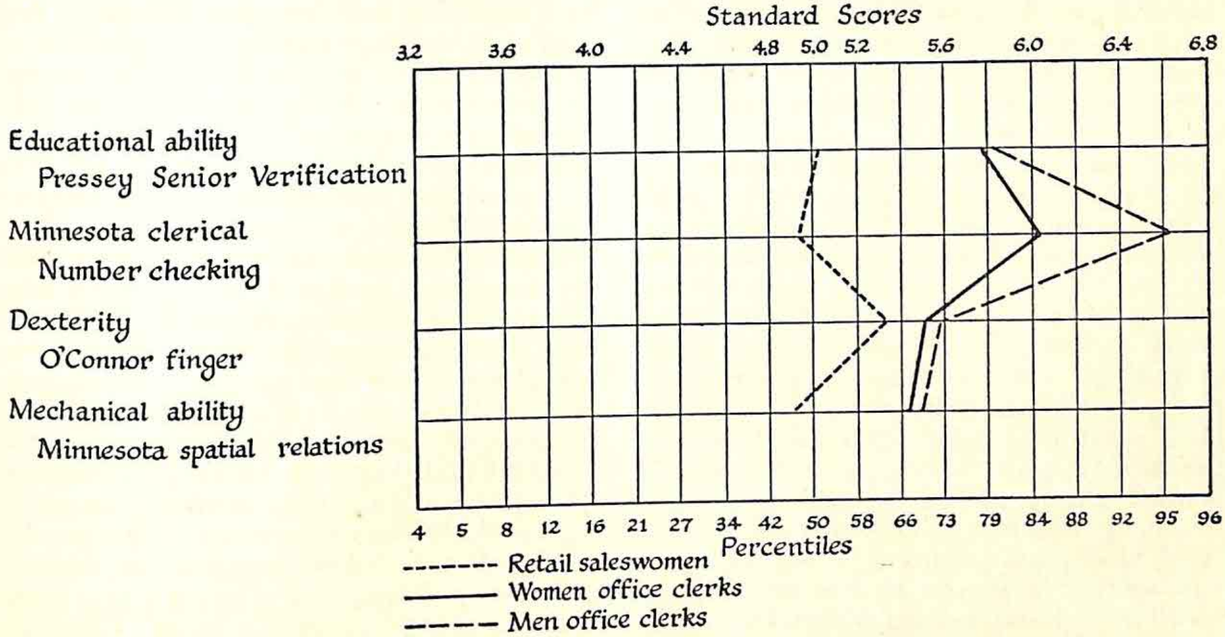


Figure 1 (V–31): *Occupational Ability Profiles*

Comparison of office clerks (male and female) and retail-sales women according to standard scores and percentile scores. Adapted from B. J. Dvorak, "Differential Occupational Ability Patterns," *Bulletins of the Employment Stabilization Research Institute,* 1935, 3, No. 8, combining Figure 1, page 12, and Figure 8, page 16, by permission of the publisher, University of Minnesota Press, Minneapolis, Minnesota.

such essential clerical abilities as concentration, mental alertness, and accuracy. Such tests are known as clerical aptitude tests to distinguish them from clerical achievement tests which measure proficiency on the job. Dudycha (1939) has shown that clerical aptitude tests measure dependability in addition to the qualities mentioned above. Manual dexterity is not usually sampled by them (the Detroit Clerical Aptitude Examination is an exception here), although most office managers regard it as important, a reflection of the problem of "face" validity which is frequently met. It has been suggested, for example, that the O'Connor dexterity measures be used for this purpose, but Andrew (1937) reports that these tests were not useful in this connection. Subsequent experience corroborates this statement: Hay (1943) found in developing a test battery for machine bookkeepers that hand, arm, and finger dexterities gave low correlations with his criterion and that mental abilities were so important that one could ignore dexterities. The minimum endowments of hand and finger dexterity necessary for typing and other clerical positions seem to be a common possession.

Lennon and Baxter (1945) have studied from a different approach the problem of what clerical aptitude tests measure. They attempted to determine what specific aspects of clerical work were related to scores on tests most commonly used in employee selection. The subjects were clerical workers in simple jobs in a large government agency and had no supervisory responsibilities. Data were reported for approximately 250 clerks, typists, and stenographers doing routine work, all of whom had previously taken a Learning Abilities Test (a local adaptation of the Army Alpha) and a clerical aptitude test developed by the authors. Those constituting the top and bottom quartiles were selected for study. The immediate supervisors of these employees filled out a check list of clerical duties and demands for each of these subjects. Then the chi-square test was used by these two authors to evaluate the significance of the difference between the high and low scores in terms of each statement on the check list. From approximately ninety statements, 12 were found to be significantly related to test scores, which indicated the specific aspects of clerical work that were predictable by means of these, and presumably similar, tests (such as understanding of work, quantity and speed of work, performance of multiple tasks). The data also pointed to the need for separate tests in typing, shorthand, statistics, grammar, and spelling for positions where such abilities were required. As might be expected, the clerical aptitude test predicted none of the personality factors thought to be important on such jobs.

Anderson (1943) in his review of clerical tests that appeared between 1929 and 1942 found signs of progress. Since 1929 more attention has been paid to reliability of such tests, and correlations now approach accepted standards; norms are now based on larger population samples than previously; but the favorite criteria are still supervisory ratings or grades of students from business courses. However, many of these validity coefficients do not meet the standards of statistical significance. It is to be emphasized that even the most popular commercially available tests are not equally sound. Often those who utilize such tests are too easily satisfied with "face" validity and generalized norms, not understanding that the presence of these two factors is not all that is needed to establish the validity of any aptitude test. The experience of the Personnel Research Section of the Adjutant General's Office (1947) is most illuminating here. A General Clerical Abilities test of six parts was developed for use in selection and placement of War Department personnel. Validity studies were made involving fifty employee groups. The criteria used were supervisory ratings and civil service position grades. The most important finding was the marked variability in the validity of the test, both as revealed between installations and between one clerical occupation and another within the same installation—further evidence that any particular test may have high validity for one occupation and low for another, despite the similarity of job titles or description of positions. Each business organization, therefore, which utilizes published tests needs to work out its own norms and check its own validity coefficients in terms of its own criteria of success. Published norms can only serve as guideposts.

MINNESOTA CLERICAL TEST

The Minnesota Vocational Test for Clerical Workers is the best known and most widely used of the clerical aptitude tests. It consists of number and name checking, and takes 15 minutes to administer. It measures clerical aptitude in terms of observation, comparison, rapid discrimination of small differences, adjustment to a new situation, and attention to a problem. Criteria for its validation were personal history ratings (a combined score worked out from experience, education, and commercial training). Test scores correlated about 0.65 with these factors. Satisfactory reliabilities have been reported. Both number and name checking, in part, measure the same abilities but their combination yields maximum validity. Andrew (1934) reported that the test was related to other clerical tests, but measured a specific ability independent of spatial, academic, and dexterity ratings. The test differentiates the rapid and average from the slow typist better than do academic ability tests. The test also distinguishes employed clerical workers from workers in general; unemployed clerical workers get significantly lower scores than do employed. A sex difference, women scoring higher than men, is consistently found from young age levels (fifth to twelfth grades) to adults. Schneidler and Paterson (1942) have suggested three tentative hypotheses to explain this phenomenon. The sex difference disappears, however, when comparisons are restricted to a given proficiency level or type of clerical work. Neither training nor experience exerts a large practice effect, although Hay and Blakemore (1943), as well as Klugman (1944), have reported score gains as a result of training and experience. Chase and Darley (1934) find a decline in scores as age increases.

Andrew (1937) suggests the Minnesota Clerical Test be used in conjunction with an intelligence test. Best prediction of clerical success was secured by weighting the clerical score 1.42 and the intelligence score by 1.00. Other workers report the use of intelligence tests as offering important supplementary data to clerical tests, especially where nonroutine jobs are involved. Bird (1931) found that discharged clerks scored lower in intelligence; Copeland (1936) reported that the Minnesota Clerical Test and the Otis Intelligence Examination supplemented each other in the prediction of occupational success.

USES CLERICAL TEST BATTERY

The United States Employment Service in its Occupational Research Program made extensive studies of clerical workers where objective criteria were available. Consequently these studies are chiefly concerned with personnel in routine jobs, such as card punch machine operators, coding clerks, calculator and adding machine operators. Specific aptitude test batteries were worked out for all such specific groups by means of the Wherry-Doolittle Test Selection Method. The possibility of combining all these separate batteries into one was then explored. A total of 562 machine-clerical workers' scores on tests, such as the Minnesota number and name checking, arithmetic, letter-digit substitution and others, were checked against production records or work samples. From the final test battery an average validity coefficient of 0.42 was found for this group; the test scores well differentiated the top, middle, and bottom thirds of workers grouped according to their production records. One can thus deal with constellations or "families" of occupations rather than with specific ones only.

Ghiselli (1942) has pointed out, however, that such an extensive test battery as this is unnecessary for most business purposes. Using the same clerical groups reported by Bellows for USES and comparing the validity coefficients on this battery against those for the Minnesota Clerical Test alone, the forecasting efficiency of the battery was markedly superior to that of the Minnesota test in only one occupational group out of eight.

THE PSYCHOLOGICAL CORPORATION'S GENERAL CLERICAL TEST

This test which was issued in 1944, is intended for general and differential use with all types of clerical workers. The content of the test is not as narrowly confined as that of the Minnesota. The over-all score may be interpreted as an index of general clerical aptitude, which correlates highly with general intelligence test scores. Three "section scores" cover routine clerical aptitude, proficiency in mathematical, and verbal facility, one or more of which might have significance for certain jobs. In addition, nine part scores are available which could be utilized in the selection of personnel for highly specialized clerical tasks. Norms are provided as guideposts to the user, but it is stressed that the most useful norms in any given case are local norms which the user compiles from his own test data.

S.R.A. CLERICAL APTITUDES TEST

The staff of Richardson, Bellows, Henry and Company (1948) has issued a test in this field that incorporates the extensive data from Army testing during World War II. The test is known as the Science Research Associates Clerical Aptitudes Test. The test samples the three independent factors found to be predictive of clerical success: office vocabulary, arithmetic, and checking. In an unpublished study conducted for the Bureau of the Census where similar problems were employed, it was found that the test differentiated not only various efficiency levels of clerical workers but also distinguished one supervisory level from another. Such a test need not therefore be limited to use with clerical workers. Local norms are again to be recommended; maximum validity will result as each user develops, upon repeated trials in his own situation, the best set of mathematical weights for the three subscores.

UNITED-N.O.M.A. BUSINESS ENTRANCE TESTS

These tests, formerly referred to as the National Clerical Abilities Tests, represent to date the most ambitious clerical aptitude and proficiency testing program of its kind. The tests have been prepared by the Joint Committee on Tests representing the National Office Management Association and the United Business Education Association. A permanent testing program was launched in 1937. The purpose of such a test program is fourfold: (1) to measure qualifications of business school graduates applying for office positions, comparable to College Board examinations; (2) to improve testing in business schools; (3) to locate student abilities and aptitudes; and (4) to assist employers in more accurate appraisal of clerical applicants. The test battery consists of the following examinations:

General information test (mental alertness, memory, etc.)
Fundamentals test (spelling, arithmetic, etc.)
Personality rating schedule
Specific skill tests for:
  stenography, typewriting, bookkeeping, machine calculation, machine transcription, and filing

The tests are advanced as "vocational ability tests" and their content is based on what employers have stated they wanted as regards abilities of clerical applicants. Statistical data, particularly those concerning reliability and validity, are sketchy; only "estimates" by experts are available. (Recent publicity about these tests, however, speaks of high reliability and validity, without offering supporting evidence.) No trustworthy norms have been published. Jackman (1938) reports that even the skill tests vary somewhat in comprehensiveness and proved adequacy. Various ones of these tests are weighted differently statistically; the interested reader may ascertain how this is done from the report by Crissy and Wantman (1942). "Experts" estimate the importance of each of the tests for a specific field and also establish critical scores for each of the skill tests. The criterion used for such critical scores is a subjective one: the minimum acceptable performance as new, skilled workers "at intermediate office levels"—whatever that may mean (*Report of Joint Committee, United-N.O.M.A.,* Bulletin No. 5). Outside criteria, it is claimed, have been unavailable.

The results of the first two years of testing were not high and few candidates could be certified. This was expected in the light of the generally poor training still prevalent in most commercial schools. The authors of the tests look forward, however, to the time when commercial school curricula will be improved to meet a large demand for student certification by means of these tests. Currently, it is felt, that by means of these tests the employer is furnished a more or less unbiased judgment as to the applicant's abilities and skills. Meanwhile it is to be hoped that reliability and validity of these tests will be experimentally and conclusively verified as is the usual procedure with psychological tests.

TESTS IN PRACTICE

Tests can be recombined often to advantage for clerical job measurement. Davidson (1938) illustrated this point when he studied the Thurstone, Minnesota, O'Rourke, and Bureau VI tests to determine the most valid sections of each for use in the selection of employees for a particular company. Correlations computed between these test scores for ninety-four clerical employees against supervisors' estimates of promotability ranged from 0.08 to 0.81. Parts of this test battery, then, differed radically in predictability. Verbal material was found to be consistently good; arithmetical problems and proverbs were low. After eliminating the "deadwood," a correlation in the neighborhood of 0.75 with supervisors' estimates was obtained. This illustration shows how a company can use ready-made tests in developing its selection devices. It further illustrates the point that particular standardized tests may be impractical for certain business offices since the tests have been developed under different conditions and with different populations from those found in the office.

OTHER TESTS

Supplementing clerical aptitude examinations are many tests of shorthand and typewriting of the aptitude and achievement type. Most business people ignore these, preferring to "test" an applicant by asking him to take dictation from current correspondence. Different problems are presented to different candidates or "catchy" or "trick" letters are prepared for dictation in the belief that tests must be difficult. Selection on this basis is one reason for the high turnover found among office workers.

One of the best of such standardized tests is the Seashore-Bennett Stenographic Proficiency Test (Psychological Corporation, 1946), which is a work-sample test on standard phonograph records. It is comprised of a standard set of five letters dealing with real business situations which are dictated uniformly to all applicants. The letters vary in length; dictation speeds vary from slow to rapid. The examination is intended for use along with a suitable intelligence and comprehensive clerical aptitude examination. A tentative scoring system, based on mailability as the criterion, has been established for the guidance of the user. Preliminary tryouts with several hundred stenographers have shown that applicants are not disturbed by the absence of a real speaker. Others have likewise found this true: Koran (1941) reports the Chicago Park District uses phonograph recordings and that the Buffalo Municipal Civil Service Commission is experimenting with radio broadcasting for the section of its standardized stenographic test requiring dictation.

Richardson and Pedersen (1947) have developed tests for the three basic skills of a stenographer. These three tests, collectively designated as the S.R.A. Program of Stenographic Skills, provide measures of language, dictation, and typing skills. The tests were designed both for improved selection and placement of stenographic workers as well as for the evaluation of student skills in commercial courses. An alternate form of the typing test is provided to mitigate against practice effects if retests are desirable; a virtue of the dictation test is that a measure of both speed and accuracy is obtainable. The dictation test also combines features of both the traditional performance (via phonograph recordings) and paper-and-pencil type tests.

Jurgensen (1942) has developed a selection test for industrial typists that does not put all the emphasis on speed and accuracy. The Sacramento City Civil Service Board has instituted certain time-saving procedures in typewriting and stenographic tests, among which is a completion-type stenographic dictation examination to measure stenographic knowledge rather than verbal memory alone (Mohrhardt and Everett, 1942). Manson (1944) has discussed the lack of attention paid to the pretest practice periods of typists and clerks; the failure to develop and use systematic practice habits before a stenographic test may be a contributing factor to the onset of emotional blocks on the part of the typist.

Mention should be made of various specialized work-sample types of clerical tests that have recently been discussed in the literature. Koran (1941) developed three performance tests for graphotype-addressograph operators, tabulating machine operators, and duplicating machine operators. Hay (1943) developed a test battery for machine bookkeepers after a job analysis showed such work involved five operations and eighteen motions. He obtained a multiple correlation of 0.70 on thirty-nine workers with the Alpha Number Series, Fryer Name Finding, and Minnesota Number Checking. Gottsdanker (1943) made a study of aptitude for crank-driven machine calculation. From a preliminary trial of nine paper-and-pencil tests, a final battery of three was selected by the Wherry-Doolittle Test Selection ("Shrinkage") Method and a multiple correlation of 0.57 is reported with the criterion. Reference has already been made to the specialized skills tests of the USES.

PERSONALITY FACTORS

Surprisingly enough, in contrast to sales personnel, the personality characteristics of successful clerical employees have been neglected. Most measurement in this field has been confined to personality traits easily quantified. The study of application blanks, for example, has been ignored. There are now beginning to appear scattered reports dealing with personality characteristics of clerical workers, a few of which might well be summarized here.

Dodge (1940) administered a personality inventory to female clerical workers of four business firms. From this he was able to identify thirty-nine items that could be scored to give a "clerical personality score," which correlated reasonably well with efficiency ratings. Prediction was best, however, with clerks on routine jobs. Dodge reported in 1942 that good clerical workers tended to be opposed to the typical personality profile of successful sales personnel; good clerks, for example, were less ready to accept responsibility, less self-sufficient, more dependent. Bills (1944) is in agreement here: the less aggressive person is the more permanent office employee. Wadsworth (1941) reports considerable success in California with the Humm-Wadsworth Temperament Scale by which means he is able to detect extreme cases of temperamental difficulties. He claims that less than one in ten individuals hired have proved unsatisfactory since 1934, largely as the result of this personality testing.

Cardall (1942) has developed a test for primary business interests from which six occupational patterns emerge: accountancy, collections and adjustments, junior clerical, sales-office, sales-store, stenography-filing. He suggests that a particular employment manager might use such a device on beginning office workers where the present scarcity of beginners places more emphasis on allocation rather than on selection.

Schneidman (1943) reports an interesting use of the Rorschach test before an experimental interview board which was considering applicants for the position of deputy clerk in Los Angeles. Such techniques are useful with higher level, supervisory personnel. From another occupational category, Balinsky (1944) reports the acceptable use, as far as management was concerned, of the Rorschach test in selecting one of two applicants for the job of plant superintendent in a machineshop. The personnel director made his selection on the basis of personality descriptions derived from the Rorschach cards. After eight months' employment, the Rorschach interpretation was checked against the shop record of the individual and showed marked agreement.

TRAINING

Two research sources provide significant suggestions on training methods for clerical workers.

BUSINESS SCHOOL PRACTICE

Results from the first United-N.O.M.A. testing program showed the need for radical revisions in business school practices. Before the construction of these tests, a committee for the Eastern Commercial Teachers Association (1937, 1938) made a survey of current high school business practices. A surprisingly large percentage of high schools had no final test, either in stenography or in typewriting. Most testing consisted of short speed spurts. It was shown that little correlation existed between ability to do the typing tasks of the daily business-office regime and speed in copying from straight printed material of the usual school tests. The survey showed that school practices need to be changed so that students can become more proficient in working steadily with either shorthand or typing; that proficiency testing in schools should be revised to conform to this plan of instruction; that more time should be devoted to machine transcribing. In this connection Deemer's study (1944), comparing the Script and the Gregg shorthand systems, is of significance. High school students were studied over a two-year period. The study is also interesting because of its departure from the usually employed stenographic accuracy and speed criteria. Seven specific criteria were developed to measure shorthand efficiency either in Script or

Gregg. Most of these measures favored Script which was, for example, superior to the Gregg system in both accuracy and speed of transcription. A follow-up study, consisting of a questionnaire answered by students who had graduated and ratings for employer use, disclosed little difference between the two systems. The latter part of the study showed, however, that more than 50 per cent of the students involved were unable to find jobs where they could use shorthand. Shorthand training ought possibly to be limited to fewer students; certainly the amount of school funds and student effort wasted on shorthand is great. The suggestion was made that students train for work on computation, dictation, and other machines rather than exclusively on shorthand.

Too few business schools make any attempt to select their entrants; it is too often assumed that everyone has the requisite aptitudes for such training. For such a purpose the Bennett Stenographic Aptitude Test is recommended for use in testing ability to acquire the skills of shorthand and typewriting. It has given better results than those predicted by standard intelligence or manipulative tests.

Barrett (1946) reports that many Hunter College students who enroll in typewriting and stenography courses, in the hope of increasing their employability, later find that they have to drop out because of their lack of necessary aptitudes. With the aid of psychological tests, it was found that good versus poor typists and stenographers could be successfully differentiated by means of scores on the Minnesota Clerical Test, certain subscores from the MacQuarrie Test for Mechanical Ability, and the Turse Shorthand Aptitude Test.

A bulletin of the Minnesota Employment Stabilization Research Institute (Eriksen, et al., 1934) contains helpful suggestions which stand in contrast to the usual business school practice. Sixty-three men and 16 women were trained 21 weeks in shorthand, typewriting, bookkeeping, business English, and office practice. The Dvorak Simplified Keyboard was used for beginners in typing. It was found that age, amount of formal schooling, and previous occupational experience were unimportant in the training of typing skill. (The previous occupations of the five most successful students were: window cleaner, steamfitter, janitor, housepainter, warehouse worker.)

A few examples from this bulletin will illustrate what has been termed the "direct approach" in training. The shorthand course contained no word drills or theory. From the beginning all material was in sentence form; later, paragraphs were introduced. All material increased in difficulty as training progressed. Speed was stressed from the beginning. All tests were dictated in sentence or letter form. The students read their own shorthand notes; practically no engraved notes were used. Students took letters from different instructors to simulate work conditions on jobs, whereas, in the usual stenography course, the same individual gives all the dictation. The typewriting course was similarly treated. Nonsense syllables and straight copy work were not emphasized. "The quick brown fox jumped over the lazy dogs" may have all the alphabet letters, but it is in no way relevant to business.

ENGLISH USAGE

A training requisite for clerical work is adequate English, especially in stenography. O'Rourke (1934) has pointed out that extra time might be more profitably spent in improving English than in securing high speeds in dictation. Transcribing dictation at the rate of 96 words per minute (the Civil Service passing score) was found by O'Rourke to be greater than that of 55 per cent of stenographers in industry. Ability to use English correctly was far less satisfactory. Industrial concerns, by and large, are satisfied with the speed with which the stenographer can take dictation but dissatisfied with the English used. Ninety-six words per minute is as high as is needed by most business companies. Above this standard, then, and allowing for some necessary overlearning, schools should be concerned with other training problems than emphasizing higher stenographic speed.

Brown and Smith (1937) worked out a program for the integration of English and typewriting in six junior high schools of Tacoma, Washington. All pupils were trained on the Dvorak Simplified Keyboard and used a specially prepared text (now available from the American Book Company). The typing group which had no English classes made greater gains in the learning and use of capitalization and spelling than did the English-study group; on tests of punctuation and English usage, the latter group made greater gains than the former. Evidence is presented, however, that the total gains in the English abilities tested, summated arithmetically in terms of scale points (a questionable procedure at best) were greater for the typing group than for the English-study group. It is thus practicable, at any rate, to include a typing course in the first year of junior high schools concomitantly with an English class.

Bean (1946) has developed an English usage test for clerks, typists, and stenographers for the Louisiana Department of State Civil Service, which represents an improvement on many such tests. The examination, after elimination of nonsignificant questions by an item analysis and a still further reduction of items as a result of consultation with experts in the field of business practice, gave moderately high correlations with both the Purdue Placement Test in English and with the A.C.E. intelligence examination. Scores were not significantly related to age, therefore individuals just out of school are not given any special advantage.

TRAINING THE TYPIST

The literature on training suggestions for typists is voluminous; the most substantial part of these reports has been summarized in Blackstone and Smith's book on typewriting instruction (1936). Barton (1926) found the whole method superior to the part in teaching the keyboard to beginners. Entwisle (1928) has questioned the value of rhythm in typewriting; Harding (1930, 1937) finds it conducive to efficient typing. Long (1929) showed the skip-around method of learning the keyboard was best. Jenson (1931), studying a small group of late elementary school pupils, found that both jazz and dirge phonograph music had a serious distracting effect on both speed and accuracy of typewriting. Davis (1938) reports a slight superiority in output with the noiseless typewriter as against the usual model. Green (1940) has made a study of typewriting derived from an analysis of weekly speed tests given to high school students during the first two semesters of work: he reported a steady improvement from week to week, despite the fact that short plateaus are reached during the second semester. Considerable research with typist training and retraining has been done with the simplified keyboard.

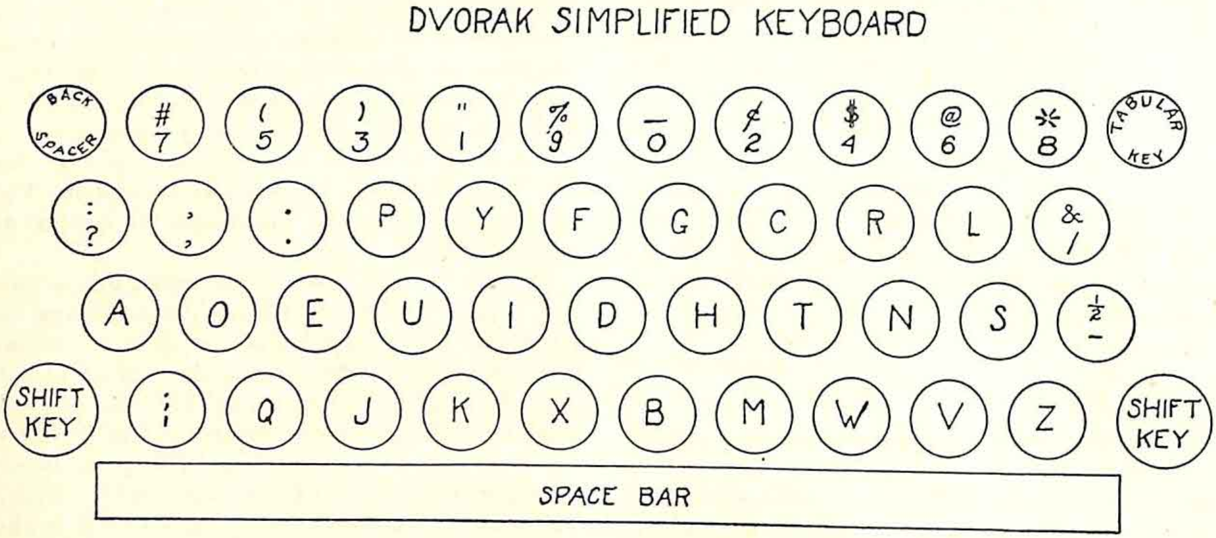
DVORAK-DEALEY SIMPLIFIED KEYBOARD

Several simplified keyboards have been devised, but the Dvorak-Dealey version (shown in comparison with the standard keyboard in Figure 2) is usually conceded to be the best. It has also had the most research done on it. There is much evidence to show that it is superior to the standard keyboard for both speed of training and permanent efficiency. But to date few machines have been sold, in part because of the lack of interest exhibited by most business schools and the opposition of typewriter manufacturing companies.

The Carnegie Foundation for the Advancement of Teaching assigned two grants in 1933–34 to the University of Washington to investigate the causes and remedies for the low average typist performance. The study revealed that the source of the difficulties with most typists lay in the faulty arrangement of the standard keyboard, an arrangement planned for the "hunt and peck," two-fingered operation. The arrangement meets the mechanical limitations of typists of 1873.

Prior to the war, publicity was given this simplified keyboard when the University of Chicago reported that fifth to eighth grade pupils in its Laboratory School learned typewriting about twice as fast with the simplified as compared with the standard keyboard (Merrick, 1941). Dvorak claims that university students can attain a rate of fifty words per minute with his keyboard after one semester's study. The standard keyboard usually takes several weeks longer. Finger loads have been considered in the organization of the keyboard so that the weakest fingers do the least work, one factor accounting for the decrease in learning time. Misspellings are also decreased. Davis (1935) studied the progress of students in typewriting at a Tacoma high school. Those learning on the simplified keyboard attained net rates in one to two semesters that students using the standard attained in from three to four semesters. Less fatigue accumulates in using the simplified keyboard because of the tenfold reduction of the awkward interplay between adjacent fingers. The slow and successive taps by the same finger have been decreased fourfold (Dvorak, Merrick, et. al., 1936). Such factors explain why the so-called "typewriter demons" found so prevalent when the standard keyboard is used occur far less often with the simplified keyboard. A convenient summary of such research and simplified keyboard benefits is presented by Dvorak (1943).

In 1944, the Navy Department, in an attempt to improve the efficiency level of its typists, conducted an experiment with fourteen standard keyboard typists who were retrained on the Dvorak simplified keyboard. After an average of 83 hours of simplified keyboard training, these typists improved their net typing speed from 32.9 net words per minute on the standard to 57.1 net words per minute on the simplified keyboard, an average improvement of 74 per cent in typing speed. Testimonials from these retrained typists indicated they were less fatigued at the end of the day, better able to keep abreast of daily typing loads, made fewer errors, and were better satisfied with the simplified key-



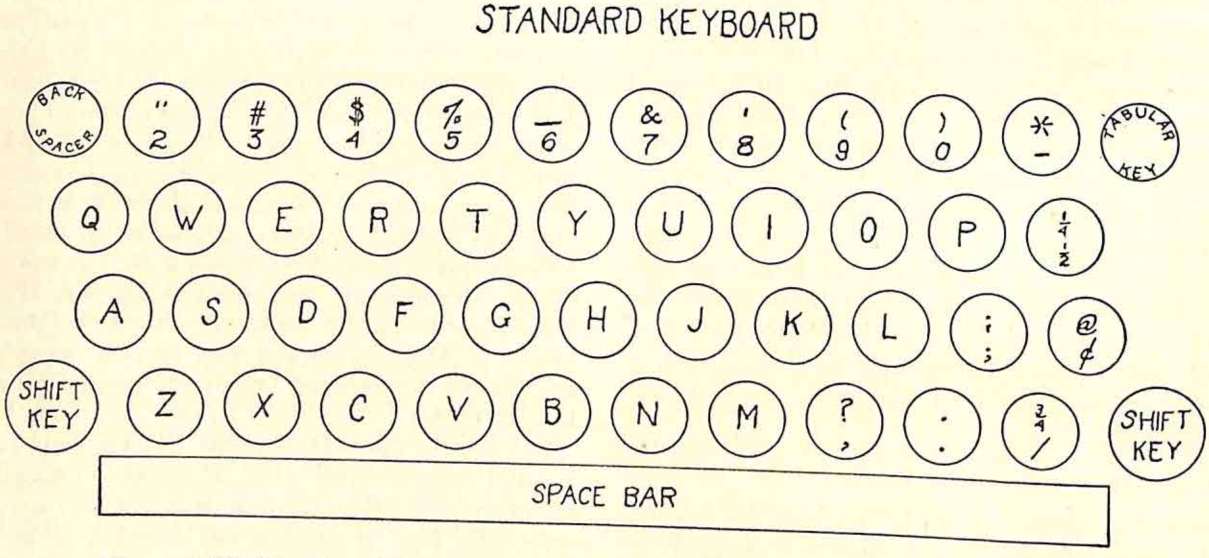


Figure 2 (V–31): *Dvorak-Dealey "Simplified" Keyboard and Standard Keyboard Compared*

The Dvorak-Dealy "Simplified" Typewriter Keyboard was copyrighted and patented, 1932, by August Dvorak and W. L. Dealey. From Dvorak-Merrick-Dealey-Ford, *Typewriting Behavior* (New York: American Book Company, 1936). Used by permission of the publishers.

board; letters from supervisors also evidenced an improvement. A comparison between typists retrained on the simplified and those given "refresher" courses under Navy auspices on the standard (eighteen operators involved in the latter group) showed that the simplified keyboard retrainees achieved three times the improvement in net typing speed, in about half the time for training and at about a third of the cost needed by the standard keyboard retrainees. To be efficient and profitable, such retraining or conversion programs on the simplified keyboard should be planned to provide a minimum of 100 hours of such training (Report of Training Section, Navy Department, 1944).

OFFICE PRACTICE

The office psychologist is concerned with numerous specialized tasks, a few of which will be mentioned here by way of illustration. He may equate jobs so that salaries are fair and equitable according to the efficiency of employees. Kingsbury's report (1933) on office positions at the University of Chicago is a good example of this work. Each job was rated for seven factors; maximum and minimum salaries were determined for each position. Salaries were based on worker competency rather than on length of service.

Sinkey (1938) has reported on the installation of work simplification procedures in a New York bank. Realizing that bank methods were honeycombed with precedent, executives had departmental flow charts constructed to show progress of work through the office. Such charts were based on the Gilbreth and Mogensen methods. Many improvements suggested by employees were incorporated into these flow charts, resulting in a considerable economy of motion.

Mogensen (1938) has worked out process chart symbols to present motion study data of office practice in condensed form. The aim here was not to increase speed of operations but to find the best way to work (usually the easiest). Tangible savings have resulted from such analyses.

Output indices of clerical workers are valuable as checks against supervisory estimates of a worker's efficiency. These can be used for salary computations, for rating an individual's promotability, and as objective criteria against which to validate selection tests. This field is practically unworked. Mitchell (1927) has given detailed instructions on how to measure output objectively for positions of typists, mimeograph and multigraph operators, and other clerical workers.

On a less technical level, McGehee and Owen (1940) have reported that the introduction of two short authorized rest pauses during the work day of 16 comptometer operators resulted in an increase in the speed of work as well as a decrease in the number of unauthorized rest pauses. Moreover the total work time was not reduced. Staples (1942) cites the good subjective effects on workers in the accounting department of the Philadelphia Bell Telephone Company attendant upon the use of four daily 30-minute musical broadcast programs.

READING REFERENCES

Andrew, D. M., and D. G. Paterson, "Measured Characteristics of Clerical Workers." *Bulletins of the Employment Stabilization Research Institute, University of Minnesota,* 1934, 3, No. 1.

Blackstone, E. G., and S. L. Smith, *Improvement of Instruction in Typewriting.* New York: Prentice-Hall, Inc., 1936.

Dvorak, B. J., "Differential Occupational Ability Patterns." *Bulletins of the Employment Stabilization Research Institute, University of Minnesota,* 1935, **3**, No. 8.

Eastern Commercial Teachers Association, C. A. Read (ed.), *Tenth Yearbook: Measures for Vocational Ability in the Field of Business Education.* Philadelphia: The Association, 1937, pp. 3–292.

———, Raymond C. Goodfellow (ed.), *Eleventh Yearbook: Modernizing Business Education.* Philadelphia: The Association, 1938, pp. 309–330.

Eriksen, E. G., et al., "Demonstration of Individualized Training Methods for Modern Office Workers." *Bulletins of the Employment Stabilization Research Institute, University of Minnesota,* 1934, **3**, No. 2.

Navy Department, Training Section, Departmental Services, Division of Shore Establishments and Civilian Personnel, *Report No. I. A Practical Experiment in Simplified Keyboard Retraining; Report No. II. A Comparison of Typist Improvement from Training on the Standard Keyboard and Retraining on the Simplified Keyboard.* Washington, D.C.: The Department, 1944.

32. MECHANICAL PERSONNEL

THOMAS W. HARRELL, PH.D.

Associate Professor, University of Illinois. Formerly: Personnel Research Section, AGO, Department of the Army, and Classification, AAF.

INTRODUCTION

In this age almost everyone is interested in mechanics because he has to operate and repair some kind of machine. Almost everyone is expected to be able to run an automobile. Almost everyone is expected to be able to do some kind of mechanical repair, for example, to be able to change an automobile tire.

As a field of personnel psychology, it is necessary to limit "Mechanical Personnel" sharply. It is fairly easy to separate from the field of mechanical personnel, the vocational fields treated by other authors in this chapter, i.e., clerical, sales, technical, police and military personnel, and civil servants, although the mechanical field touches on all of them more or less. The difference between a clerk's work and that of a mechanic is well marked in extreme cases, but there may be overlapping in the work of the clerk who runs a bookkeeping machine, and that of the mechanic who keeps books. Clerks, however, are omitted from "mechanical personnel," even when they operate machines as their primary duty. Saleswork is easy to separate from mechanical personnel even though certain salesmen have to understand, operate, or repair the machines which they sell. All sales occupations are excluded. The field of Technical Personnel runs directly into the higher skills of mechanical personnel so that the dividing line is more difficult to draw. Certain technical personnel invent new machines and plan improvements on machine design and operation. To avoid overlap with technical personnel it is necessary to omit from mechanical personnel the occupations of anyone whose skill is primarily in the invention of machines.

Mechanical personnel are found among military, police, and other civil servants. But a division of labor has to be made. It would be particularly inconvenient to omit all mention of military occupations. The Army and Navy, by providing large numbers of subjects for studies of mechanical personnel have provided psychologists with a field day for the investigation of the subject. However, those military occupations are omitted that have no civilian counterparts.

The term "mechanical personnel" is consequently limited to those people engaged in occupations that are central to the operation and maintenance of machinery.

After limiting the scope of mechanical personnel by omitting those vocational fields where the emphasis is otherwise than on machine operation or repair, there is also a residue of jobs, such as wrappers and packers, that are customarily tested by the same dexterity and mechanical tests used for machine operators. These are classified as mechanical personnel, though mechanical in a different sense from machine operators and repairmen.

THE NATURE OF ABILITY TO PERFORM MECHANICAL WORK

FACTOR ANALYSES

The nature of ability to perform mechanical work has been studied in several ways, including casual observation, job analysis, and comparison of test scores with success on the job. Perhaps the most revealing studies are those using the experimental and statistical methods of factor analysis. The space relations factor has been found central to successful tests of complicated mechanical work in each of four investigations (Harrell, 1940; Staff, Division of Occupational Analysis, 1945; Wittenborn, 1945). Tests which have been found to measure this space factor by at least two of the investigators are the Minnesota Paper Form Board and Minnesota Assembly Test.

Only one factor analysis has been found which shows a criterion of work performance definitely related to a clearly identified factor. This was the criterion of a shop product in the Minnesota investigation which was present in the spatial visualization factor (Wittenborn, 1945).

The factor analyses of mechanical tests have also uniformly shown one or two dexterity factors. A finger dexterity factor was found to be separate from manual dexterity (Staff, Division of Occupational Analysis, 1945). These dexterity factors are measured by tests that have been found more successful with routine than with complicated mechanical work.

Perceptual factors have been found in studies of mechanical ability tests, but their nature is not yet clear.

Other factors have been found in studies of mechanical tests but those mentioned above appear to be central to mechanical work. There are also numerical and intelligence factors, but they are required in many other kinds of work as well as in mechanical occupations.

APPLICATION OF FACTOR ANALYSIS

A comprehensive test battery based upon the results of factor analysis has been developed by the United States Employment Service (Dvorak, 1947). This General Aptitude Test Battery measures ten factors with fifteen tests. Mechanical work is prominent among the occupations predicted, although the battery is not limited to this. Mechanical repair and metal machining require a pattern of abilities of at lowest average in intelligence, numerical ability, spatial ability, and only finger dexterity better than the lowest one-fourth of the general population. Heavy metal structural work, plumbing, and related work require numerical ability, spatial ability, and manual dexterity better than the lowest one-fourth of the general population.

SELECTION

The selection of men for mechanical work is often made on the basis of their experience in the same or related work. The use of the interview and application blank for determining amount and quality of experience in mechanical work is not essentially different from the use of these methods in other vocational fields. There have been practically no published studies showing the effectiveness of the interview and application blank in choosing mechanical workers but these two methods are by far the two most common procedures of selection. They are used a great deal more than aptitude tests. In the paragraphs that follow summaries will be given of a number of studies correlating various tests with success at mechanical work. Even should the refinement of tests result in greater effectiveness than at present the evaluation of experience promises to continue to be the most popular method of choosing mechanical workers.

The tests used for selection of mechanical workers for the most part have been tests of aptitude or ability. Interest and personality tests, although possessing possibilities have been little used in this vocational field. A few studies have shown positive correlations between visual tests and success on routine mechanical jobs.

ROUTINE MECHANICAL OCCUPATIONS

MACHINE OPERATORS. Experimental selection studies have been reported, in terms of correlation coefficients, for operators of sewing machines, punch presses, paper mill converting machines and the like. There have been several studies reported correlating tests with the success of sewing machine operators. The Blum Sewing Machine Test correlated 0.42 with success for 50 operators but a followup of the test results for 17 entering employees failed to show useful prediction (Long and Lawshe, 1947). The Number Comparison Items of the Minnesota Vocational Test for Clerical Workers correlated 0.40 and 0.50, and the Minnesota Paper Form Board (Likert and Quasha, Series AA) correlated 0.31 and 0.48 with the efficiency of two groups of sewing machine operators (Stead and Shartle, 1940). A punch press test yielded significant correlations with ratings of speed, accuracy, and safety for 25 punch press operators (Long and Lawshe, 1947). Including a new procedure, the Minnesota Rate of Manipulation correlated 0.66 with ratings of 60 converting machine operators in a paper mill (Jurgensen, 1943).

ASSEMBLERS. A hand precision test correlated 0.49 with production of 33 plug-socket assemblers (Tiffin and Greenly, 1939). In the same report finger dexterity, hand precison, visual acuity, and color blindness correlated 0.60 with the ratings of forty-two radio assemblers. A spatial perception test correlated 0.56 with a criterion of thirty-one action assemblers (Link, 1919). The Thurstone Identical Forms test correlated 0.66 with ratings for one hundred and forty assemblymen (McMurry and Johnson, 1945).

PACKERS. The success of thirty merchandise packers correlated 0.49 with Series AA of the Minnesota Paper Form Board (Stead and Shartle, 1940). The Minnesota Paper Form Board and the Turning part of the Minnesota Rate of Manipulation Test correlated significantly with ratings for a small number of pharmaceutical inspector-packers (Long and Lawshe, 1947).

INSPECTORS. A cancellation test correlated 0.63 with the production of 51 shell inspectors, but with thirty Paper Shot Shell inspectors the correlation was not significant (Link, 1919). With the larger group of fifty-one inspectors Link also found two other tests to correlate significantly, Card Sorting 0.58, and Number Group Checking 0.72. Thurstone's Identical Forms test correlated 0.71 with ratings of sixty-nine inspectors (McMurry and Johnson, 1945).

PACKAGE WRAPPERS. A study by Ghiselli of forty-two package wrappers yielded insignificant correlations between ratings and the Minnesota Rate of Manipulation Test, and the O'Connor Finger Dexterity (Long and Lawshe, 1947).

WORK INVOLVING TWO HANDS. Evans has described a pegboard which was correlated with fifteen samples of jobs involving the use of both hands (Long and Lawshe, 1947). The median validity coefficient was 0.34. This presumably is a test, or the forerunner of a test used by the United States Employment Service.

SUMMARY OF SELECTION RESULTS FOR ROUTINE MECHANICAL WORK. The results are not impressive. A number of investigators have obtained positive, even significantly positive, correlations of validity. Usually the number of cases were small, and the number of variables relatively large. The total number of tests is large. A frequent interpretation made from these data and from intercorrelation results is that manual ability is highly specific to the particular industrial job. Additional consideration will be given to the nature of manual ability as it is related to mechanical ability. At this point, however, it is well to point out another alternative explanation. If one uses a sufficient number of variables and few enough cases it is likely that high correlation coefficients will result. Aside from the investigations of the United States Employment Service no instance was found of cross-validation. No two investigations found the same test to correlate significantly with a criterion for the same job.

SKILLED MECHANICAL OCCUPATIONS

The dividing line between routine and skilled mechanical occupations is not an easy one to draw. In general machine operators are classified in the routine group and machine repair men in the skilled group. In addition, metal workers are classified usually in the skilled mechanical group.

MACHINE REPAIRMEN. Grades in auto mechanics courses correlated 0.42 with the O'Connor Wiggly Block (National Industrial Conference Board, 1941).

Three studies which gave the results of correlations with the various courses in training for aircraft mechanics have been reported. In one of these studies various combinations of tests correlated significantly with each of four out of five courses (Jacobsen, 1943). Grades in the course in Aircraft Engines had no significant correlation with any test. There were approximately 80 cases. The Aircraft Instruments course learner ratings correlated 0.64 with a combination of scores on the Pressey Senior Verifying Test, the Bennett Mechanical Comprehension Test, the O'Connor Tweezer Dexterity Test, and the Frandsen Arithmetic Prognosis Test. Learner ratings in each of three other courses, Aero Repair Mechanics, Machine Shop and Aircraft Electricity had significant multiple correlations with a combination of test scores.

Two studies of Army airplane mechanics have been reported which give a cross-validation on several paper and pencil tests (Harrell and Faubion, 1940, 1941). The only tests to stand up in the two groups of 84 and 105 men were two of Thurstone's tests, Surface Development and to a lesser extent, Mechanical Movements. Correlations with the Mechanical Drafting and Blueprint Reading course were the highest of any of the subcourses. For the Surface Development Test the correlations with this course were 0.54 and 0.50; for the Mechanical Movements Test, 0.39 and 0.26.

An adaptation of the Stenquist Mechanical Assembly Test correlated significantly with ratings of excellence for forty-five loom fixers; and even more significantly with ratings of mechanical ability of forty carding machine fixers (Harrell, 1937). As a follow-up, 91 additional loom fixers were given the Minnesota Mechanical Assembly along with the other Minnesota Mechanical Ability Tests and a number of other mechanical aptitude tests (Harrell, 1940). No validity correlations were significant which was probably due to the fact that several plants were tested involving different sets of raters whose ratings were not comparable.

MACHINISTS AND TOOLMAKERS. With ratings for 105 toolmaking apprentices a combination of the following yielded a correlation of 0.48: Geometrical Construction Test, Circle Making, Errors in Number Checking, and Age (Pond, 1927). In the same investigation subtests similar to those on Army Alpha did not produce significant results. With another group of 83 toolmaker apprentices each of the following tests correlated between 0.32 and 0.39 with rankings: Kent-Shakow Industrial Form Board, O'Connor Wiggly Block, Scovill Classification Test, and MacQuarrie Test of Mechanical Ability.

A combination of the Stenquist Mechanical Assembly Test, the Cube Test, and a Form Board correlated significantly with rankings of 35 apprentice machinists and machine tool specialists (Link, 1919).

Machine Tool Operators. Two tests were found to correlate significantly with rankings of 66 machine tool operators (Bennett and Fear, 1943). These were the Mechanical Comprehension Test, Form AA, 0.64 and Hand-Tool Dexterity, 0.46.

Shop Grades. In addition to training records mentioned above there have been several studies reporting significantly positive correlations between shop grades and tests. The criterion for the widely used Minnesota Mechanical Ability Test was such a one, namely a grade based on the quality of a shop product for seventh and eighth grade boys. For 100 cases the tests correlating highest with this criterion were: Minnesota Assembly, 0.55; Minnesota Paper Form Board, 0.52; and Minnesota Spatial Relations, 0.53 (Paterson *et al.,* 1930). Another study of shop success showed a significant correlation of 0.49 with the Revised Series AA of the Minnesota Paper Form Board (National Industrial Conference Board, 1941). In the same study, which was made by the United States Employment Service, the same test was reported as correlating 0.49 with Mechanical Drawing Grades. O'Connor's Wiggly Block correlated 0.57 with marks of 109 high school shop students (Remmers and Schell, 1933).

Welders. Thurstone's Identical Forms test correlated 0.64 with ratings of 97 welders and the Mechanical Comprehension test correlated 0.66 with ratings of 140 welders (McMurry and Johnson, 1945).

SUMMARY OF SELECTION RESULTS FOR SKILLED MECHANICAL OCCUPATIONS

The prediction of success in skilled mechanical occupations is somewhat more substantial than is the case with routine mechanical work. There are, however, few studies that confirm the validity of a particular test for a particular occupation among two populations. One exception is the Surface Development Test which predicts the training records of a Mechanical Drafting and Blueprint Reading Course taken by airplane mechanics. As a rule the studies of skilled mechanical workers have been in terms of correlations with training records rather than objective measures of success on the job. It seems likely that several popular tests will predict the training success of mechanical workers in the maintenance field. The Minnesota Form Board has been a successful test in predicting training records. O'Connor's Wiggly Block, in spite of low reliability, has shown significant validity in several samples. Thurstone's Identical Forms Test has been tried in only a few studies but has shown considerable validity. One explanation for divergent results in different populations for the same occupation is the amount of prior selection that has taken place.

TRAINING

The purpose of training has been said to be to overcome the effects of selection. Individual differences between one man and another frequently lead to profound differences in learning time and in the performance of industrial work. The purpose of training is to bring everyone, however chosen for a job, up to a satisfactory level.

Some kind of training program is provided for all mechanical as well as other jobs. At the one extreme there is a training specialist whose primary task is to teach the job to new employees. At the other extreme is the situation in which the new employee learns his job from watching the operator next to him. The most customary method for mechanical workers, particularly the routine ones, is to receive training from the supervisor. Some of the most skilled mechanics, especially in the metal trades, are trained by the apprentice method.

The desirability of systematic training is so self-evident as to have elicited few elaborations of its precise effects. The effects of training as well as of selection are most dramatic when there is a sudden need for the quick production of many new products. Such a need is most pronounced during wartime. During World War II the training of mechanics for civilian industry was tremendously important. The need for rapid training is shown in the fact that 7½ million workers were being added to the total labor force of the United States while 10 million workers who joined the Armed Services were replaced (War Manpower Commission, 1945). The bulk of additions were in manufacturing where mechanical work is prominent. That production per man-hour increased in spite of the large number of workers new to industry was no doubt due in part to the motivation of patriotism, but at least equally important was the tremendous program of training within industry.

Training within industry, by common usage shortened to TWI, was carried out by the conference method with almost 2 million supervisors (War Manpower Commission, 1945). There were four separate courses or sets of conferences, Job Instruction Training (JIT), Job Methods Training (JMT), Job Relations Training (JRT) and Program Development (War Manpower Commission, 1945). Each of these courses was made up of 10 hours of conferences. The conferences were conducted with groups of 10, or less, to assure universal participation. An outline of the main points of each course was presented by the conference leader and each point was put across by the supervisors giving cases from their work experience to illustrate them.

The training within industry programs were tremendously effective. Plants throughout the United States that used the programs have reported on specific improvements. These reports were tabulated in terms of whether production, for example, increased as much as 25 per cent after the training programs had been used. A summary of results adapted from a War Manpower Commission Report (1945) follows:

Result	*Percentage of Plants Reporting Results of 25 per cent and Over*
Production increased	63
Training time reduced	92
Manpower saved	73
Scrap loss reduced	53
Grievances reduced	96

The coordination of training and selection is a desirable objective in mechanical work as in the other occupational fields. Definition of training standards set objectives for selection. Measurement of training results provide criteria for selection procedures as well as an indication as to how well or how poorly training has succeeded.

READING REFERENCES

Dvorak, B. J., "The New USES General Aptitude Test Battery." *Journal of Applied Psychology,* 1947, **31,** 372–6.

Harrell, T. W., "A Factor Analysis of Mechanical Ability Tests." *Psychometrika,* 1940, **5,** 17–33.

Long, W. F., and C. H. Lawshe, "The Effective Use of Manipulative Tests in Industry." *Psychological Bulletin,* 1947, **44,** 130–148.

Paterson, D. G.; R. M. Elliott; et al., *Minnesota Mechanical Ability Tests.* Minneapolis: University of Minnesota Press, 1930.

War Manpower Commission, *The Training within Industry Report.* Washington, D.C.: Government Printing Office, 1945.

Wittenborn, J. R., "Mechanical Ability, Its Nature and Measurement: I. An Analysis of the Variables Employed in the Preliminary Minnesota Experiment." *Educational and Psychological Measurement,* 1945, **5,** 241–260.

33. FIELD SALESMEN

MARION A. BILLS, PH.D.

Assistant Secretary, Aetna Life Affiliated Companies, Hartford, Connecticut; Diplomate in Industrial Psychology of The American Board of Examiners in Professional Psychology. Formerly: Assistant Director, Bureau of Personnel Research, Carnegie Institute of Technology; Consultant, Life Insurance Sales Research Bureau, New York; Chairman, Testing Committee, Life Office Management Association.

Selection of field sales personnel is a separate problem from that of selection in the clerical, professional, or production fields. The tools used, however, are the same. They include: (1) personal history, (2) references, (3) interview, and (4) tests. The emphasis in the selection of salesmen is on personal history and the interview.

PERSONAL HISTORY

The psychologist has turned the application blank into a scientific instrument by studying in detail the relation of the various personal history items to success in selling. Certain items currently found on application blanks have proved of no value in predicting success. Some have proved of value for certain types of selling and of no value for others. A few appear to be of definite value in each of a number of independent fields. Among the last are the following.

AGE

Investigations carried on in widely different industries have shown an effect of age on the percentage of men entering the field who succeed. The Life

Agency Management Association, the Aetna Life Affiliated Companies, and the Tremco Manufacturing Company have found definite age ranges in which a larger percentage of men succeed in selling than at other ages.

EDUCATION

Some studies have shown a relation between education and success in selling; in other studies no such relation has been apparent. In general, education has been found important in those types of selling in which the article sold is such that understanding or explaining it requires a high degree of mental alertness. One factor has proved to have predictive value: the person who has completed the section of school he has entered is likely to be more successful than the one who leaves before completing that section. For example, in several studies it has been found that the best education for the salesman is to have completed college. The next best is to have completed high school. The third best is to have completed eighth grade and not to have entered high school; while the poorest of acceptable salesmen come from those who completed only two or three years of college, or one, two, or three years of high school. This criterion for obvious reasons became invalid during the war period.

MARITAL STATUS

Almost all investigations have shown that the married man is more likely to be successful in selling than the unmarried. However, this advantage is only apparent with men who are twenty-five or over.

DEPENDENTS

There is general agreement among investigators that to have from one to three dependents is a desirable factor in predicting the success of salesmen and that to have no dependents is undesirable. Beyond three dependents there is disagreement as to prediction: some studies have indicated that the more dependents, the better the chance of success, while others have indicated that dependents beyond three were less desirable than one to three.

CLUBS

In general, to belong to no clubs is disadvantageous for the salesman, but statistics have shown that for the salesman of certain products, belonging to one or two clubs is more advantageous than belonging to a larger number.

OCCUPATIONS

Almost all studies have indicated that previous occupation, the length of service, and remuneration form discriminating factors for predicting future success in selling. In general, to have had previous selling experience seems to point toward success. To be employed is more advantageous than to be unemployed, and if unemployed, the shorter the time the better. Length of time on previous jobs should have been long enough to indicate that one is stable, but not so long that one becomes set in one's way. The previous salary which most nearly approaches the range that can be hoped for on the new job seems to be the one most predictive of success. In the employment of salesmen under twenty-five, occupation of the father has some predictive value, the more favorable occupations being those most closely allied to selling. The young man whose parent is a salesman or a businessman is more likely to succeed in selling than one whose parent is a professional man or a day laborer.

FINANCIAL STATUS

Paralleling in its predictive value the previous jobs held is the present financial status of the applicant. Financial status has been judged by items on the application blank giving information of insurance carried, amount of debts, actual net worth, and minimum current living expenses required. Investigations of the Life Agency Management Association indicate for life insurance salesman that the larger the amount of insurance carried the more advantageous for successful selling. Investigations of the Aetna Life Affiliated Companies for casualty salesmen and of the Tremco Manufacturing Company for technical salesmen show a relation between age and amount of insurance. Up to twenty-five years of age carrying insurance is a positive factor in predicting sales ability, while beyond that age it is only positive if more than $5,000. The explanation may be that if under twenty-five years of age one has not had time to set a comprehensive insurance program, while if over twenty-five and able to buy life insurance at all, a small amount may indicate a poor financial situation.

INTERVIEWING

Very few salesmen are hired without at least one interview with the prospective employer. While much has been written concerning the theory and technique of the interview it has lent itself to little analytical study. A few positive statements can be made from research, however: (1) most interviews are too short; (2) the interviewer does too much of the talking, the interviewee too little; (3) a directing guide to an interview or a rating of the applicant at the time of the interview or immediately following gives better results than the undirected interview or the remembered rating; (4) there should be more than one interview, and interviews spaced a few days apart are better than all interviews being given on the same day.

PURPOSE

The interview should not be used primarily to secure factual information which can be obtained equally well from the application blank or from references. It should be used to obtain attitudes which are more important and which often are impossible to understand in any way but by direct conversation. It is the interviewer's task to cause (by skillful conversation) attitudes to be displayed by the candidate. These attitudes are dependent, not only upon what the individual says and does, but also upon the whole situation of which he is a part. The interviewer has to perceive these attitudes and to form a judgment upon the basis of what he perceives. One might find that a person worked overtime in his last job, information which could be obtained from records; but if the person's willingness or unwillingness to have done this work is revealed the interview will have got at something which could have been obtained by no other tool.

INTERVIEWING BLANK

Guides to the interview may be elaborately prepared blanks, on which the details of the past history of the applicant are checked and rechecked, or they may be simple rating sheets on which the interviewer gives his general estimate of the interviewee. It is important to have them sufficiently detailed so that the interviewer can recall the person in future discussions of him with other interviewers or in comparing him with other applicants for the same job. It is also important that the interviewer's rating be so arranged on the interviewing blank that it will give material for a scientific evaluation of the interview when the success or failure of the candidate can be ascertained from definite performance.

SELLING THE JOB

The interview can be used as the means of explaining and selling the job to the applicant. Here there is one ironclad rule: paint the picture darker than it really is, not brighter.

TESTING

APTITUDE

Since mental alertness or intelligence tests have proved so useful as selection aids for many occupations, it is disappointing to find that they are only of limited value in the selection of salesmen. Other factors outweigh what is measured by pencil-and-paper tests of intelligence.

It has been found that for certain technical selling there is a minimum intelligence score for success; and for certain types of routine selling there is a maximum score above which the turnover becomes so high that it is better not to hire. However, various studies in the last decade have brought out the fact that for many sales jobs requiring ordinary mental alertness there is little relationship between intelligence test scores and success in selling.

INTERESTS

For at least twenty years the interests of an individual have been used as an aid in the selection of salesmen. A number of questionnaires are used for this purpose. The Strong Vocational Interest Blank is one widely used. It is a list of four hundred items for which the person indicates his likes or dislikes. From a comparison of these likes and dislikes with those of people who have been successful in a given field, the individual's sphere of interest may be ascertained. Investigations have shown a positive relationship between the correspondence of a person's interest with those of successful people in a given field and his own future success in that field. A prediction of success or failure better than chance can be made for persons applying for jobs as life insurance salesmen. A study over the last eight years in the Aetna Life Affiliated Companies indicate the same to be true for casualty insurance salesmen. In general, a man scoring high has three times the chance of success that a man has scoring low.

PERSONALITY

Personality adjustment has always been recognized as one of the outstanding factors in the success of salesmen. Various tests have been tried with the idea of selecting men who are dominant, extroverted, self-sufficient, and emotionally well adjusted. All of these tests have consisted of series of questions to be answered and all of them are "bluffable." What makes bluffing difficult to detect is that most of it is unconscious rather than conscious.

These questionnaires function most successfully when disguised or when forming a part of a larger testing program so that the individual taking the test is less self-conscious, as in the aptitude index for life insurance salesmen as developed by the Life Agency Management Association. This index is a combination of an application blank and a personality questionnaire; the first part is obviously a personality rating blank and the last two parts are much less obviously so. This is one of the programs for the selection of salesmen which has been most carefully worked out and for which conservative and well substantiated claims are made.

APPLICATION

In setting up a program for the selection of field salesmen the following is recommended: (1) a fairly complete application blank which can be definitely scored for predictive value of success and failure; (2) an interest analysis which, not only can predict better than chance success or failure but also may prove an opening wedge to an intensive interview; (3) an interview which is outlined, directed, and recorded, and which forms a reference for future consideration and sells the job to the individual; and (4) a combination of all of these factors into a composite whole from which the selection is made.

READING REFERENCES

Bills, M. A., "Relation of Scores in Strong's Interest Analysis Blanks to Success in Selling Casualty Insurance." *Journal of Applied Psychology*, 1938, **22**, 97–104.

Kenagy, H. G., and C. S. Yoakum, *Selecting and Training of Salesmen*. New York: McGraw-Hill Book Company, Inc., 1925.

Life Agency Management Association, *Selection of Agents*. Hartford: The Association, 1937; *Measuring Aptitude for Life Insurance Selling*, 1938; *How Well Does the Aptitude Index Work?*, 1941.

Lovett, R. F., and M. W. Richardson, "Selecting Sales Personnel: The Significance of Various Types of Test Material." *Personnel Journal*, 1934, **12**, 248–253.

Ohman, O. A., "Research on the Selection of Salesmen at the Tremco Manufacturing Co.," New York: American Management Association, *Marketing Series*, 1941, No. 45, 11–17.

Oldfield, R. C., *The Psychology of the Interview*. London: Methuen and Co., Ltd.

Phoenix Mutual Life Insurance Company, *Selecting the Successful Salesman*. Hartford: The Company, 1937.

Strong, E. K., "Interests and Sales Ability." *Personnel Journal*, 1934, **13**, 204–216.

34. RETAIL SALES PERSONNEL

KATE LEWIS, M.A.

Personnel Director, Namm's, Inc., Brooklyn, New York. Formerly: Head of the Department of Philosophy, Rockford College; Instructor in Psychology, Maxwell Training School; Lecturer in Retail Personnel Administration, Hunter College and Long Island University.

Retailing, which in 1941 had become the third largest enterprise in the United States, exceeded only by manufacturing and agriculture, now has a 74 million dollar business. It employs 6 million people. Economic conditions have changed old ideas of the department store as a haven for the untrained or superannuated worker from the professions, trades, or the household. In 1938, Congress passed the George Dean Act which appropriated a million and a half dollars for the training of people in the retail field. To do this legislators must have been somewhat certain of the need for such training and of a good return on the taxpayers' money.

The 6 million workers are of different calibre from those of twenty years ago; they are superior in education, in ability, and in effectiveness. In the small town their selection probably is a very personal one on the part of the owner of a store. He trains them day by day. Competition that the large department store faces causes the responsible executives to perform the selection and training functions continuously and as scientifically as is practicable. This is the main reason for the improvement in department store sales personnel.

PSYCHOLOGICAL PROBLEMS

Merchants have too much at stake to neglect the study of the human beings who are so necessary to their success. They are sensitive to such remarks as: "I'll never go there again; the girls are so hateful and they don't know a thing." That is the kind of advertising that no store owner wants if he can help it. The layman does not always realize that his complaints of rudeness or ignorances are utilized in the selection and training of workers. Firing for a single mistake is gone, but correction of workers' errors is part of a long-range plan for instruction.

LABOR MARKET

General business conditions have a direct effect on the department store's selection processes. In an employee's market the experienced stenographer, bookkeeper, or telephone operator will apply for work in offices where she envisions shorter hours and better pay than are offered in the department store. These same persons may be in the department store sales force during an employer's market.

Employment managers in department stores have seen their work change from investigating fitness for positions open and advice about positions elsewhere, in one era, to turning down people who would take any job in another, and finally, in a third era, to advertising jobs in their most attractive light. These are differences actually found in department store personnel administration over a twelve-year period. Wars make a tremendous difference in the labor market for retail selection. Very often at the end of a war the merchant finds himself with personnel whom he would not normally employ.

SELECTION

A specific knowledge of many lines of work is required in department store personnel administration. Jobs run from heavy unskilled labor through mechanical technicians, and include all grades of clerical to professional work. And in no other kind of work are they more closely interlocked. There is not much sense in picking successful sales people if the cashier who finishes the transaction enrages the customer enough to make her throw down the purchase and walk out. It is not good business to have the satisfied customer wait in the store's restaurant a starvation period for service because a dietician is unable to direct the waitresses in an efficient manner. Nor is there any profit in having a customer receive merchandise that is not correct. A furniture salesman may have done a magnificent job of understanding the problems of a newly married couple in selling them a chair needing a little alteration, but if the cabinet-maker does not handle the alteration satisfactorily the future account may be lost. If copywriters and artists cannot do work equal in quality to that of the store's competitors the store is not going to attract customers. A variety of interrelated occupations is involved in department store selling.

Selection has to be made efficiently and profitably to meet these conditions. Turnover must be kept down; it is admittedly inefficient and expensive. Moreover, in unionized stores, it can become very time consuming for the personnel department.

TEACHING FUNCTION

Selection involves instruction of the employment interviewer as to how to turn away hundreds from his door. To do this diplomatically he must become something of a teacher. The usual applicant thinks of retail selling as an easy job. Although he may recognize the need of training in the trades or a profession, he feels he can always do selling. It is the employment interviewer's business to judge whether he can be trained to sell immediately, or whether he can reach that goal after serving a period in stock or clerical divisions, or whether he belongs in the business at all.

In a department store, there is not a sharp dividing line between clerical work and sales work. The sales force has at times a great deal of clerical work to do and there are a great many clerical workers who sell. A store is a service organization. Even the operators of billing machines and the Convenient Payment accountants must realize the relationship of their work to the customer. So the employment manager or interviewer must take a long-range view of persons applying for sales positions. He must teach them what they can or cannot do in a department store.

COUNSELING FUNCTION

The employment interviewer also becomes something of a guidance counselor. Good human material may be wasted if he selects for a blind-alley job in his own company someone he knows is promising for another field. Such selfish procedure leads to dissatisfaction on the part of the employee which in time can cause bad morale throughout a department.

The selection of workers has to be made, not only for today but for tomorrow. The majority may remain contented performing routine skills for years. There needs to be, however, in selection a consideration of those who can some day become executives. Good morale is maintained in a department store through promotion of the qualified from within the store, and bad morale is in part the result of promotion of those unqualified. While the employment manager may ostensibly be picking salespeople, if he does his job right he is picking the future buyers, store managers, and other executives.

From such consideration in selection the applicant gets a feeling of security at the very beginning of his work-life in the store. He feels that he has been chosen because of ability and capacity, not because of favoritism or politics. Hence he wants to live up to the reputation he has established through selection, and he expects the store to use his abilities as indicated in this procedure.

TRAINING

Training a salesperson before he goes on the job means less waste and more immediate productivity. If a salesperson does not know store routine before beginning work, he is confused and causes irritations to customers and fellow-workers, and loss of sales. Trial-and-error learning on the job of salespeople is a thing of the past. Its cost in inefficiency is so great that training programs for new employees are now usual among large department stores.

The feeling of security that is engendered in the selection process is enlarged by the training program. Training is planned for the person who has never worked in a department store. The pattern for good habits is laid in order that bad habits may not creep in so easily. If a mistake is made after the applicant has gone to work, the system of correction already established is there to break a bad habit in formation. With such an objective procedure the salesperson's reaction is more apt to be one of learning or apology, rather than of resentment. If there is resentment, training has broken down because of lack of guidance in the selection and training procedure.

STORE TRAINING

The organized training program of the modern store is a far cry from the groping aids given to salespeople of two decades ago. Research workers in retailing were then just beginning to learn the futility of giving the eighth-grade salespeople books to read to improve themselves (Charters, 1925). They found they had to put salespeople into the classroom. The high school and college graduates in department stores today have their duties analyzed for them. Their production is better because of the research that management has carried on.

But more interesting than this intelligent organization of the routine training program is the realization of how far training can and must go. The "Training within Industry" programs of the war years are now being applied to the retail field. In the handling of union grievances the training of the supervisory force in human relations becomes a recognized necessity. It is hardly necessary to point out that the understanding of individual and group psychology is essential in these two fields.

Department store executives not only have recognized that thorough initial training means greater productivity, but also have seen that continuous training will keep their salespeople out of ruts. Salespeople in department stores are made conscious of the necessary learning process at every stage of development or change in their work.

PSYCHOLOGICAL INFLUENCES

Retail selling of all kinds presents a practically untouched field to the psychologist. There was a period when employment departments indulged in the buying of many tests and in the giving of tests in great numbers to applicants. Industry was using tests and finding them valuable. Retailing followed suit, but it used tools that it had not made its own. Industry was doing slow, intelligent research on tests under the guidance of qualified psychologists. Such unfortunately was not the way in retailing. Personnel people realized that their methods of selection needed improvement. But 90 per cent of them were not qualified to use the new tools and the other 10 per cent found that lack of time for research or that more seemingly important work were obstacles to any satisfactory discoveries in this field.

Reference in the preceding paragraph is to tests for the selection of that large body of people in the retail store on whom so much depends—sales personnel. This seemingly simple job performed so often by the untrained person or by many different persons is waiting for intensive and broad psychological research. Research will need to consider many more variables than it finds in industry. The act of selling may seem the same to the customer, but it is a different act every time it is performed, not only in separate establishments, but within the one store itself. Herein lies the broad scope for research work.

The distributive services in the country perform their work in diverse ways. A manufacturing salesman sells in one way, an insurance salesman in another; a grocery salesman in another, a department store salesman in still another. They are all salesmen, but success in one field does not necessarily presage success in all. The department store salesperson, in fact, may be a success in one store, but fail in another; moreover, the salesperson in a single store may do well in one department but be unable to carry the load in another.

Retail sales work probably is foremost among the occupations requiring a specific sales psychology. But the retail salesperson is untrained in psychological theory of personal contact and probably cannot be trained in it for practical reasons. He gains a practical knowledge of the psychological problems of selling by watching customers' reactions, considering what approaches are most effective, restraining egoistic retorts, and teaching the customer to want the article. In the past, personnel executives of department stores became practical psychologists in much the same manner through dealing with applicants for positions. Other methods of training are now available.

GRADUATE TRAINING

The graduate training for department store executives in the retail training schools includes various amounts of psychology. A great majority of the graduates of these schools enter the merchandising division where, after a period of internship, they begin their climb in business. But some elect the personnel division. All of these students have had a few courses in psychology, but they will have absorbed their psychology from a merchandising point of view more than from a study of individual and social differences and conflicts.

A number of personnel men and women have received training in personnel administration and vocational guidance and have higher degrees from large universities. They have had instruction in tests and measurement and in interviewing, and have achieved the personnel point of view. But few with this training enter department stores; the hours of work are better and psychological practice is better established according to the personnel point of view in industry. However, the mercantile field offers them an absorbing laboratory of human problems.

The door to the laboratory is opening slowly but cautiously and steadily to the trained psychologist and technician. The mistakes of twenty-five years ago, made by untrained people who spoke glibly of correlations and rejection of applicants below a raw score which had never been validated, are not apt to be repeated. Tests and psychologists must prove themselves in this field. The present quiet research that goes on in a few places is a good sign.

SELECTION

The selection process has not lagged behind training in psychological applications as much as it might appear. Studies made of the validity and reliability of the interview have contributed to the interviewing technique in use in department stores. There have been two reasons why the employment interviewer did not use psychological measurements in his selection of salespeople: (1) psychology has not presented him with a test that measures sales ability and (2) with practical interviewing techniques he has a more than even chance of not burdening the production force with poor salespeople. Then, he usually has had an enormous labor market for selection of workers.

The interviewer in the retail field has to employ people for a small community. If he does not select a particular person for a selling position he may select him for another type of work. Psychological tests did not seem to contribute materially to his task when the time element for his work was taken into consideration. There have been changes in this thinking.

During World War II the labor market was practically nonexistent. Standards for selling dropped. Customers took what they could get. Business was good. But the war did stop. Merchants are conscious now of operating costs. Salaries are up. No store, either union or nonunion, really wants to cut salaries. How then can this high standard be maintained and a profitable business still be run? The trade journals advise their readers that better selection must be made; more analytical training and more attention to personal problems of the workers must be given. Is testing one help toward answering these problems?

Department store personnel staffs are in general exceedingly conscientious; they are eager to improve their own selection techniques. Where management sustains research, psychologists and technicians receive enthusiastic support from the personnel division.

USE OF TESTS

The present period of psychological research in retail selling is devoted primarily, though not entirely, to exploration. Very interesting work is being done to help in counseling on behavior and low-production problems. Some preemployment testing has been done and needs time for its validation. On the whole the work seems to be setting a steady pace, not just a sudden burst of speed at one place as has happened in the past.

HISTORICAL POINTS

In 1935 an experiment was begun by the Worker Analysis Section of the Division of Standards and Research, United States Employment Service, in which three department stores cooperated. This experiment was about the only important one that was made. In 1938 the National Retail Dry Goods Association, through its personnel division, made a survey of testing in retailing and found that stores covering less than one eightieth of the total retail population used tests in their personnel administration. These included tests of achievement of a homemade variety.

By *1941* the same National Retail Dry Goods Association checked the test practices of thirty nine member stores, which roughly represents volume and sections of the country. About two thirds of them used some form of testing for part or all classes of preemployment applicants. Basically the tests were of the informal type. Ten stores reported the use of Standard Psychological Tests, such as the Thurstone Typing Test, the Otis Mental Alertness Test, the Johnson O'Conner Workroom Sample 16. Emphasis was on nonselling activities. An increase in interest in testing was noted, and it was also recognized that people with psychological training were needed. The feeling was that "aptitude testing" had proved less practical and successful in the retail field than it had in industry. Critical analysis would have shown that industry's success was made possible by its different approach to finding out about tests. Money had been set aside for psychologists and research and experimentation. Time was given. This approach was to be used a little later in the retail field.

The Worker Analysis Section of the United States Employment Service had developed an experimental battery from the work of 1935, which was released in 1939 to forty-six public employment offices. This battery was prepared particularly for use with those who apply for department store sales. By 1940 another battery was released to the employment offices for variety-store use. This recognition of different types of selling is an aspect of the work that retail management will appreciate as a step in the understanding of its peculiar problems. The bureau expected that much further work would have to be carried on.

In 1945 the Research Bureau for Retail Training, University of Pittsburgh, began a long-range research project. In cooperation with five well known department stores, they began an exploratory testing program. The objective was to find, if possible, a series of tests that would provide information helpful in the selection of retail personnel. Somewhat different batteries are being used in the different stores. The groups tested range from applicants for routine jobs to applicants for promotion to executive positions. Part of the testing is used for training; part as the basis for counseling. Results of this research have not as yet been published. These and other independent stores have found accredited tests helpful on the supervisory level; the research concerning the sales group is still far from complete.

FRINGE ATTACKS

A challenging approach to the study of the "sales personality" is by means of the interaction chronograph, but the method awaits further research for its validation. Dr. Eliot C. Chappel, of the Department of Anthropology, Harvard University, in cooperation with the Gilchrist Store in Boston, has been analyzing the successful and unsuccessful salesperson by means of this technique. In this store it seems a successful objective aid to the selection of salespeople.

Many of the salespeople employed in the United States are in the grocery stores. One large chain in Texas has been carrying on an experiment since 1945. A consultant was first used, later a psychologist took over the work. This experiment is concerned with preemployment selection. Norms for the old employees were established, using the experimental aptitude test for supermarket personnel from the Psychological Corporation. The Bell Adjustment Inventory, Wonderlic's Personnel Test, and the Minnesota Rate of Manipulation Tests were also given. The United States Employment Service permitted their Grocery Checker battery to be used. Dr. Ruth Rabson, who is conducting this experiment in addition to another on the selection of personnel for drugs, domestics, glassware, china, and infant's-wear, reports that much more work is needed before the results can be evaluated.

The individual salesperson is being studied, too, from the point of view of personality adjustment, but not necessarily in the store itself. The Board of Education in Pittsburgh, Pennsylvania, has set up a Community Vocational Testing and Counseling Center. Stores in the vicinity send their people to the center.

Research on the elusive sales makeup is also being undertaken in The New York State Institute of Applied Arts and Sciences, which was created after the war to provide special education at college level for technical employment not requiring a college degree. The curriculum in Retail Distribution is intended especially for the student who plans a career. Psychological tests are given to all these students. The student's schedule is so arranged that he spends one month in school and the next in a department store. Tests have been given in Abstract Reasoning, Verbal Reasoning, Language Usage, Clerical Accuracy, Numerical Ability, Space Relations, Mechanical Reasoning. The students will be closely followed. This, too, is another of those long-range projects that the retail field has started. The results will find a waiting public.

READING REFERENCES

National Retail Dry Goods Association, *Current Operating and Personnel Problems.* New York: The Association, 1946, pp. 91–97.

——, *The Development of Personnel Testing in Retailing.* New York: The Association, 1938.

——, *Testing as an Aid in Replacements.* New York: The Association, 1941.

Piel, G., "Your Personality Sits for Its Photo." *Nation's Business,* 1947.

35. MANAGEMENT'S USE OF PERSONNEL RESEARCH IN TRAINING

ROGER M. BELLOWS, PH.D.

Chairman, Department of Personnel Methods, School of Business Administration, Wayne University; Vice President, Richardson, Bellows, Henry and Company, Inc.; Editor, *Industrial Training Abstracts;* Diplomate in Industrial Psychology, American Board of Examiners in Professional Psychology. Formerly: Personnel Procedures Officer and Executive Officer, Personnel Research Section, AGO, Department of the Army; Lecturer, New York University; Associate Professor, University of Maryland; Supervisor, Occupational Research Program, US Employment Service; Instructor, Ohio State University.

Enlightened management is turning from hit-or-miss personnel methods to systems determined by personnel research. The principal difference between traditional methods and personnel policy controlled by personnel research is that the latter is concerned with its own "goodness." Continuing evaluative checks tell management wherein its personnel policy is succeeding, wherein it is failing. Planned training can be an integral part of a personnel research program.

Heneman (1946) makes a plea for expenditure of 1 per cent of the total pay roll for personnel research. Employees were paid more than 79 billion dollars in 1942. At 1 per cent the amount for personnel research would have been 800 million dollars. However, most experts agree that personnel research can save 5 per cent of the pay roll over a period of time, which would have been 4 billion dollars for the entire country in 1942.

Training is a function of management, continuously being performed under its direction in all occupational groups, such as the clerical, mechanical, and sales. The next topic will attempt to show the tie-ins of training with the various personnel functions of management, and later the contribution to training of personnel research and psychology will be discussed.

TRAINING TIE-INS

Training is interrelated with each area of personnel work. To fail to take account of these interrelationships is like trying to use gears that do not intermesh. To achieve the best results, training should be coordinated in practice with communications and employee relations; recruiting, selection, induction, and placement of employees; job analysis; pay roll control (job evaluation and merit rating); employee services;

safety; morale and motivation; and turnover control. These relationships will be noted briefly.

COMMUNICATIONS

Every manager and personnel director knows the role that training plays in communication, although the use of vertical and lateral channels of communication are sometimes neglected. Most of the findings of the Western Electric studies, for example, point to this need (Mayo, 1945). Both worker and management need knowledge of and about the other. Planned training may economically contribute to communication.

Management and supervision boil down to the training of others to do the job. Most of the time of top management is spent in talking to, instructing, directing, or counseling other supervisors. Whether this is done by direct conversation, written memoranda or manuals, or through another supervisor, the function and result are the same. The end result of training is change in performance of the employees so trained. It makes no difference whether the change is brought about by a formal, planned program called training, the point is that training is going on.

The real problem here is whether the job of communicating can be done more economically and effectively by a planned supervisory training program. The answer depends, of course, on the situation. Consideration may well be given to this question, as many companies may be missing an effective means of needed communication because of lack of planned training. Two examples may be considered briefly.

During the years of World War II Bill Jack of Jack and Heintz appeared to have a close personal relationship with his workers (Ward, 1946), engendered by his personal appearance in the buildings and by his frequent informal talks over the public address system. These channels of contact, for that situation, may have been both satisfactory and economical. In another firm organized channels of communication were so remiss that supervision, removed only two or three echelons from the top, knew company policies only through fire escape and grapevine channels of communication. For Jack and Heintz a formal training program for opening up communications may not have been necessary; for the unnamed firm an effectively planned training program for top, middle, and bottom management and workers might have been the most economical means of developing communications.

RECRUITING, INDUCTION

How is training related to recruiting, selection, induction, and placement of new employees and executives? Recruiting, in times of a tight labor market, is made doubly difficult for a particular firm as a result of poor communication procedures. Company personnel policy as understood (or more often misunderstood) by the worker is a general influence in the community which constitutes the labor source. If the company is unfavorably considered in the community, few of the better workers will apply. Planned training may be tied in profitably with recruiting.

The outcome of the selection and placement interview is in part based on the interviewee's knowledge of the jobs available in the company. Has the applicant had an opportunity to learn what the jobs are like so that he may state a choice? How did he get his information? Often a public relations policy entailing a close integration of planned training with community schools may be of help in the selection and placement of employees.

Induction is clearly a part of planned training. The war-developed induction training programs of certain General Motors Corporation divisions, notably G-M Truck and Coach, at Pontiac and AC Spark Plug, at Flint, could serve as models for many other companies.

JOB ANALYSIS

How is planned training related to job analysis? How does it relate to use of the products of job analysis—job descriptions and job specifications? First, job analysis requires, for its successful initiation, a considerable amount of indoctrination of supervisors. They want to know what it is, how it is done, what it yields. Their questions lead logically to the training, formal or informal, of supervisors as well as the highly specific program for training job analysts.

Then, how are job analysis products, i.e., job descriptions and specifications, to be used with profit? The answer is clear. Their use requires that the users be trained to handle them, the same as any worker is trained to handle skillfully a new tool.

Job descriptions form an economical basis of centralized training program development. Suppose you are developing a new job-methods training program. You have a job-methods committee working on the development of this training program.

You already have job descriptions. You start with job descriptions of the jobs as they now exist and refine the job elements. This is but one of the many training uses of job descriptions.

Job specifications, including minimum hiring requirements for each job and its relationship with other jobs, are useful in training employment interviewers and in induction training of new employees. Such specifications contain basic information relative to organizational relationships and promotion possibilities, items of intense interest to the newly inducted employee.

JOB EVALUATION

Then again, training is closely related to job evaluation procedures in salary and wage control. Here it has a three-way relationship: (1) indoctrination of all concerned, so that the developmental program of job evaluation will be accepted; (2) training of the committee that works out the problems; and (3) training supervisory personnel in the use of the resulting tool.

If job evaluation is to work properly it must be preceded and accompanied by a carefully planned indoctrination program. Bi-lateral participation is needed—joint worker-management endeavor—if job evaluation is to be of value (Percival and Cross, 1946). Otherwise the resulting information and tools would go unused.

The training program for use of the products of job evaluation may best be considered by reference to the topic outline of a training manual developed in a specific situation. It includes the following headings: aims of job evaluation, description of the program, how the program works, and the values of the program.

MERIT RATING

Planned training is equally important for the initiation, development, and use of a merit rating or personnel inventory system. A personnel inventory simply will not work without appropriate indoctrination of management supervision as raters and employees as ratees. This fact is abundantly clear when the several aims and outcomes of a well conceived personnel inventory system are considered, which are as follows: (1) to evaluate employees to enable personnel actions such as transfer, upgrading; (2) to sharpen up supervision: motivating supervisors to observe, describe, evaluate, and train employees; (3) to offer an additional incentive for workers; (4) to enable the ratee to know clearly how he is doing on his job, for his satisfaction and morale; (5) to enable the ratee to know his weaknesses with a view to personal development and self-improvement; and (6) to yield a check on various personnel procedures, including training methods.

Acceptance of these aims presupposes that rating data and personnel descriptions will be discussed individually with each ratee. Such discussion may usually be done most effectively by a personnel counselor; the program may best be planned by the training department.

The approximate use of this service in training is exemplified by the following excerpt from a memorandum regarding a rating system to a group of professional workers:

> A merit rating procedure has been devised to aid in evaluating in a uniform way the professional work and performance of each member of the staff. A copy of the merit evaluation form together with a manualized description of its use is attached for your information.
>
> You may wish to study the method so that you will know its purpose and on what basis supervisors judge your work. Your attention is invited to the performance items that are included, as these may be of particular interest. You will be rated only by your immediate supervisors who are thoroughly familiar with your work.
>
> After each periodic rating you will be given an opportunity to know the evaluations of your performance on each item, and to discuss them with an office executive with a view to your further professional development and improvement.

EMPLOYEE SERVICES AND ACTIVITIES

These activities require orientation training of supervisors and induction training of new employees. The purposes of the services, and how to use them, can in most situations be most economically described and their proper uses discussed by planned group training programs. If properly done, much good will can be engendered in this way; if left to chance, the expenditures for such services may not be justified. For example, Evans (1945) reproduces a stenographic record of how *not* to induct new employees. The transcript of the introductory remarks of the speaker reads: ". . . Now if any of you are still lucky enough to own automobiles, you'll find plenty of parking space out front, *but don't make the mistake of pulling into a space that has a name on it. You don't belong there, so don't get into trouble.*" Today this company uses sound motion pictures to present properly and with correct emphasis the facilities and services provided for the employees.

PERSONNEL SAFETY ADMINISTRATION

This subject is concerned with planned training. Industrial accidents cause about 17,000 fatalities each year in the United States (Heinrich, 1930). It has been estimated that about 98 per cent of all accidents are preventable, and that only about 10 per cent are attributable to physical working conditions (Heinrich, 1931). The importance of tieing-in training as a tool for prevention of accidents is clear. Accidents have multiple causes. It is known that some workers are accident prone—a great proportion of accidents happen to a few workers. Selection plays an important role. After the worker is selected and placed, training is significant. Inexperienced workers have more accidents than experienced ones, and length of time on the job is known to be inversely related to number of accidents in certain situations. Age is known to be a factor. Emotional stability and mental attitude, the latter probably somewhat susceptible to measurement and change by training, are also factors in safety.

MORALE AND MOTIVATION

These topics are closely related to training. One of the crucial steps in the introduction of an incentive system, a merit rate increase system, or any new personnel tool is the attainment of a high state of employee morale. Planned training is needed here to accomplish two needs. The first is to enable study of attitudes, beliefs, and opinions of the workers. There are several ways to get at worker attitudes: employee counseling is one, questionnaire survey is another, attitude measurement scale is still another. Whichever is used, the training department is usually the logical channel for administering the attitude-probing procedure. It will be readily agreed that any kind of attitude measurement is a special case of communication in which information flows upward from worker to management.

The second need for planned training in handling the problem of morale and motivation is concerned with the introduction and administration of methods intended to improve morale and motivation. An example is an incentive system, such as that used by the Lincoln Electric Company. Another is an employee suggestion system. Others are concerned with the announcement and introduction of changes in materials processing brought about by technological developments. Whatever the innovation, planned group training is an effective means of indoctrinating employees, which constitutes a means of communication whereby information flows downward from management to employee.

Turnover may be effectively controlled by aid of planned training. An efficient training plan may well include employee counseling with a view to controlling avoidable turnover. Such plans would enable greater use of the principle of vertical expansion, i.e., upgrading from within the organization. Companies which not only give lip service to, but actually practice vertical expansion would, it would be reasonable to suppose, have lower turnover rates.

REALISTIC BASIS OF TRAINING

Having touched on a few of the personnel activities which are related to planned training, we turn now to discuss a realistic basis of training and the need of personnel research as a foundation for the planning.

The War Department General Staff made a wrong guess in one gigantic training operation during World War II. It took out of the Army Specialized Training Program many thousands of partly trained men and placed them in Army Ground Forces training programs. This action was based on best estimates of replacement losses for establishing beachheads in Europe. The only previous experience the War Department had for making these estimates was the Anzio beachhead losses. As it turned out, fortunately, these replacement losses did not occur in Europe and millions of man-training-days were "wasted." While Army's guess was wrong, it was the best possible guess under circumstances of war.

Business and industry are not run like war. There is no reason why a firm should not know, within plus or minus 2 per cent, the number of foremen, or apprentice mechanics, it will need a year from now.

The rule here may be simply stated—train only the number of men that will be needed. The actual functioning of this rule requires that company organization, policies, and managerial statistics be effective enough to enable prediction of the number of men that will actually be needed.

Closely related to number of graduate trainees needed is the problem regarding the number of trainees who start a program and then drop out for one reason or another, and those who complete training at company expense and then resign to use their training and experience in another company. There being no indenture system whereby trainees must remain with a firm after training, the firm would do well to study, as a part of planned training, reasons for such dropouts and resignations, with a view to selecting for training those who have a high probability of remaining with the company. This problem has caused concern in several large companies, among them General Motors Corporation, The Ford Motor Company, and Chrysler Corporation.

A realistic basis for planned training is concerned, not only with numbers of trainees actually needed, but with content of training courses. This is a matter of determining "where the emphasis shall fall," as it is stated by those who develop training programs at the G-M Institute at Flint. It is not easy to determine where the emphasis shall fall. Two sources of information will aid in determining this: carefully obtained judgment of executives who know company policy, operations, and needs; and job analyses yielding detailed descriptions of the duties to be performed by graduate trainees.

METHODS FOR TRAINEE PLACEMENT AND EVALUATION

The Director of Personnel Research at Owens-Illinois Glass Company, Reign Bittner, maintains that tests reduce training time, and Edward Hay, Personnel Officer for the Pennsylvania Company, says "Tests are the one most valuable tool." In a recent survey of one hundred companies it was found that 45 used tests. Of these, 56 per cent indicated they use tests for selecting trainees and apprentices (Routt, 1946).

Whether tests and related selection devices may be used with profit in planned training depends upon the situation. The considerations of a training situation that lends itself to profitable use of trainee tests include: the number of trainees to be trained; the characteristics of the applicant group; and the length, content, and expense of the training course.

Clearly if only a few trainees are to be trained, the proper answer is, "Why bother with trainee selection tools?" If the applicant group is homogeneous, i.e., if there is adequate reason to believe that all applicants are qualified, the answer is the same—don't use tests. If the training is only of a few hours or days duration, consideration other than aptitude may be more important—if so, don't try tests. If the content of the courses is such that all of a certain class, department, or job incumbent should have the material, such as indoctrination in a new company policy, selection is not the problem, therefore tests are not helpful. If the training organization is immature and if the intended outcomes are not clearly known—don't try tests—to work on organization and planning first would be best. Profitable use of tests for selecting trainees may come later, when the company is ready for them.

Examples of the proper use of tests (the term is used here to include any systematic trainee selection device of known dependability) may be cited from intensive Army studies as yet unpublished. Trainee selection devices—interviewing and merit rating methods—were developed at the end of World War II by the Personnel Research Section of the Army specifically for selecting postwar regular Army officers for further training and development. They have implications for any personnel program involving selection of executive, administrative, and supervisory trainee personnel in business and industry.

The Army was faced with a problem of sizable proportions—the selection of men to be retained as officers in the regular postwar Army. Only a small proportion of those holding wartime commissions and wishing to remain in the Army could be retained. With a view to selecting the men for career training as fairly and economically as possible the problem was referred to the Personnel Research Section of the Army. Here was a situation which lent itself to the use of tests and other selection tools.

Out of the work of the Personnel Research Section involving "trial runs" with approximately 31,000 officers in more than fifty Army installations grew three methods of officer selection which were found to do the job better than any previous method used: the Officer Evaluation Report, the New Interview Board, and the Biographical Information Blank.

The Officer Evaluation Report is simply a carefully standardized rating scale which makes it possible to rate an officer on his performance and to compare him with other rated officers of the same grade.

The New Interview Board replaced the loose, multi-purpose traditional Army interview board. Whereas the traditional board sought, in a rather haphazard fashion, information concerning a candidate's intelligence, physical fitness, technical competence, and personality characteristics, the New Interview Board was concerned with the evaluation of the candidate's ability to get along with people in line and staff functions. Physical fitness was left for a medical examination; intellectual fitness was determined by objective tests; and the evaluation of technical competence, background information, and certain personality characteristics was the function of the objective information blank. A novel aspect of the New Interview Board was that it succeeded in what might seem to be two mutually exclusive tasks, namely, maintaining an air of easy informality during the interview and at the same time adhering carefully to standardized instructions. The special interview form used consisted of two work-sheets upon which to record observations, candidate reactions and ratings in three areas, bearing and manner, voice and language, and personality traits.

The Biographical Information Blank provided for the record of background information of a personal, educational, or occupational nature along with self-descriptions of traits, attitudes, and habits.

With these three trainee selection devices in hand, the crux of the matter was: How well will they discriminate good from poor officers? Can they be used with any assurance that they select better than other, older methods? Let us examine in detail the way in which these particular tests were tested against "real-life" officer performance. The success or failure of trainee selection methods rests squarely on the care which is taken to determine whether or not the methods are valid, that is, whether they actually do measure what they are supposed to. In this instance the Army wanted to know if good officers as distinguished from poor ones are discovered by the methods.

Approximately 15,000 officers were brought together in small groups of from fifteen to thirty officers who were well enough acquainted with each other's work to evaluate it. Each officer was requested to prepare two lists of officers in his group—those high and those low in general all-around value to the Army. From a third list each officer was asked to designate the five men in the group whom he believed to be most closely medium or middle with respect to over-all value. Only officers who fell clearly into one of the three groups—high, middle, or low—in the sense of virtually perfect agreement of their fellow officers were used for further study.

The efficiency of the various instruments will be described in terms of how well they succeeded in picking high and eliminating low officer trainees for retention. The Officer Evaluation Report was the most successful of the three new devices and did a better job of selection than did the average of the previous efficiency ratings. If each of these rating devices is considered separately as a predictor for selecting the top third of all applicants, the Officer Evaluation Report correctly selected 64.2 per cent of the high officers against 62.3 per cent of high officers picked by the previous efficiency report, and erroneously picked only 5.6 per cent of low officers whereas 12.7 per cent slipped through on the previous efficiency report. While the average previous efficiency rating does a fairly satisfactory job, the Officer Evaluation Report appeared somewhat more efficient in discriminating at all levels of officer efficiency.

The Biographical Information Blank and the Interview did less well than either of the rating methods, but much better than the traditional Army Board and, when added to the Officer Evaluation Report to secure a combined Point Index, further improved the discrimination between high, middle, and low officers, assuring a still better retention program.

Table 1 shows the percentages of officers according to these three groups which would have been selected by a Combined Point Index of 240 (used as a cut-off score, which would have selected the top 30.9 per cent). From Table 1 it can be seen that the Officer Evaluation Report was superior to the average previous efficiency rating; each of the three proposed instruments was superior to the traditional Army Board; the Combined Point Index, based on all three instruments, was superior to the previous efficiency report (PER) and greatly superior to the traditional Army Board, and the traditional Army Board method was only slightly superior to chance. The Combined Point Index, on the other hand, doubled the number of high officers who were selected and decreased the low officers from thirty-one to only two per hundred.

Table 1: Percentage of Officers Falling in Each of Three Criterion Groups, and Correlation Coefficients, for the Several Selection Procedures

	Criterion Group			
Selection Procedure	*High*	*Middle*	*Low*	*r*
Officer Evaluation Report (OER)	64	30	6	.60
New Interview	58	25	17	.39
Biographical Information Blank	52	30	18	.35
Combined Point Index	74	24	2	.67
Previous Efficiency Report	62	25	13	.45
Traditional Army Board	45	26	29	.09
Chance	37	32	31	.00

From these studies it was concluded that the Combined Point Index will tend to select officers satisfactorily for regular Army training on the basis of past and present performance. If one considers that it costs at least $100,000 to induct, pay, train, maintain, and retire an officer selected for commissioning in the Regular Army, and if one further assumes the value of any low officer, judged among the poorest by *all* his fellow officers, to be zero (this may be too high an estimate of his worth if he decreases the morale of a division, makes decisions leading to disaster in combat, or alienates the affections of enough congressional committee members), then the adoption of the program, since it selected only two instead of thirty-one low officers, could mean a saving of $2,900,000 for every hundred officers selected.

Another example of profitable use of tests for selection of trainees involves the Army General Classification Test. The Personnel Research Section of the Army developed the first form of this test in the summer of 1940. One or the other of its four equivalent forms was given to more than 13 million army personnel. It was found early in the war that the Army General Classification Test had dependable validity for selecting trainees for officer candidate schools. For this reason a critical or "knockout" score was set at a standard score of 110 (this eliminated the lower 69 per cent of the Army population, or selected the most promising group comprising the top 31 per cent). For experimental purposes, and other reasons, a few men scoring under 110 were admitted to officer training schools. Let's see what happened to them! Table 2 shows the relation between failure in these training situations and Army General Classification Test standard scores. The data in the table are based on a study of 710 students in the Engineer, Signal Corps, and Coast Artillery officer candidate schools.

Table 2: Army General Classification Test Scores and Failure in Officer Candidate School

AGCT Score	Percentage Failing	Percentage Commissioned
125+	10	90
110-124	31	69
100-109	54	46
90-99	74	26

It can be seen from Table 2 that of those men whose standard scores on the AGCT were between 110 and 124, about 31 per cent did not make the grade. More than half of those in the group whose standard scores were below 110 did not make the grade. The saving in money and manpower effected by a critical score of 110 for admission to Officer Candidate School in the Army, one can deduce, is gigantic. The implications of these facts for planning for industrial training situations are clear.

ECONOMY IN TRAINING

To neglect several of the more important rules of economy in training is to disregard the cost. These practical training rules have been developed largely through research by the educational psychologist. The quantitative, scientific approach to the study of learning was started by Ebbinghaus in 1885. Now in 1947 there is a considerable heritage of training literature from the fields of education and psychology, much of which is applicable to industrial training situations.

Only a few of the general rules have been selected for discussion here, with special reference to those for which research evidence is available. They include: train the trainers, motivate the trainees, plan in terms of individual differences among trainees, overtrain, plan regular training intervals, and plan in terms of the reading ability of your trainees.

Train the Trainers. It would seem but common sense to insure that trainers know their subject and know how to instruct others in it. Frequently in industry, however, procedures for uniform certifying and qualifying standards are almost totally lacking. Therefore little planning is given to this important item. It may be of interest to have a look into an investigation which demonstrates the profit that may result from training the trainers.

In a study by Bavelas, reported by Maier (1946), trainers were given 8 hours of special training, which included methods for obtaining favorable social relations, procedures for increasing motivation, and teaching procedures that would tend to engender a more favorable attitude on the part of the trainee. Trainers were taught by the discussion method, not by the lecture method.

A marked increase in rate of learning by beginners was noted at once. Whereas beginners, in learning stitching operations, could learn after one week to achieve only 18 units of work before the trainers were trained, beginners could learn to do more than 27 units when the instructors were trained. Table 3 shows these results.

After 12 days of training, beginners trained by instructors who themselves had received no special instruction, produced 25 units while those learning under trained instructors produced 38 units. The results of the study further suggested that instructors

Table 3: Amount of Learning When Trainers Were Untrained and When They Were Trained (after Maier)

Days of Training	Units of Work Performed by Beginners	
	Trainers Untrained	After Trainers Had Been Trained
5	18	27
10	24	33
12	25	38

who were given more specialized training were better than those who were given general training. The moral is clear—plan for training by giving due consideration to training the trainers!

MOTIVATE THE TRAINEE. Much has been written about motives and incentives in learning. Many experiments have been performed. To boil these down we can say simply that learning is more likely to take place rapidly if the trainee is motivated, that is, if there is a goal-incentive to be achieved by learning. This goal-incentive for the trainee may be increased pay or social status, approbation by fellow-workers or supervisor, or self-satisfaction in the knowlege that he has learned.

What will likely pay off, is the rule, closely allied to the motivation problem. Learning is likely to be more rapid when the learner has knowledge of the results of his attempts to learn (Forlano, 1936). This rule is one phase of incentive or motivation.

The question arises as to how to inform the trainee of his progress so that he will be learning with knowledge of results. While this depends somewhat on the situation, achievement tests, quantitative units of production during learning, and the careful use of trainee rating devices as counseling tools are good methods to employ.

PLAN IN TERMS OF INDIVIDUAL DIFFERENCES AMONG TRAINEES

Greater differences in the ability exist in trainees than is usually evident to those who plan training programs. This statement is true even though a considerable amount of screening and selection of trainees has taken place. Test data from training situations in the Army in World War II clearly substantiate this statement. Judicious selection of training materials and methods, as well as careful sectioning of trainees in terms of background and basic skills, such as reading ability, should follow an understanding of the facts of individual differences.

OVERTRAIN. It is known both from experience and from experiment that skills and knowledges once learned are soon lost. A third or more of college sophomores are below fifth grade norms in arithmetic skills. Within two years college students lose more than two thirds of what they have learned in a course. Typists who attain 50 words per minute in training drop to forty or less within a year without retraining or practice. The US Civil Service Commission has been much concerned over the rapid deterioration of skills of stenographers who, at the time of appointment, were found by testing to be highly proficient in taking and transcribing dictation at from 90 and 120 words per minute.

What is the solution? The answer, in planning a training program, is to plan to overtrain. Simply plan to train the workers during the training period to a criterion of proficiency higher than that deemed necessary for the job. If 70 units are desired, bring trainees to a point of proficiency in which they produce 90 units.

PLAN REGULAR TRAINING INTERVALS. Economy in training may depend to a considerable extent on spacing the practice intervals (Starch, 1912). Usually spaced practice or training periods are more effective than massed training periods. The over-all reduction in learning time achieved by distributing practice has been known since the time of Ebbinghaus and has been verified by studies of many different kinds of tasks since that time.

Closely related to planning regular, spaced practice periods is the notion that what is retained after a practice period depends also upon what kind of activity follows the practice (Jenkins and Dallenbach, 1924). The optimum conditions for economy in training is highly specific, depending on the specific training situation. In planning a training course consideration must be given to these factors for maximum economy. The interested trainer may realize much profit from review of published studies (Swenson, 1941) on this influence.

PLAN IN TERMS OF THE READING ABILITY OF YOUR TRAINEES. Twenty foremen were tested recently at a well-known manufacturing firm in Detroit which employs about 275 foremen to aid in planning a foreman training program and for other purposes. A number of tests and inventories were given, including measurement of reading speed. It will be interesting to review here two extreme cases, as a striking contrast:

Mr. X, while he has had only 11 years of formal schooling, surpassed 100 per cent of the eighth grade pupils with whom his scores were compared, and has a reading speed superior to that of the average college senior.

Mr. Y, a foreman, who has graduated from elementary school, has a reading speed equivalent to that of fifth grade pupils. He surpassed only 1 per cent of the eighth grade pupils with whom he was compared.

For economy in training, and for general utilization of employees, there are three implications: (1) recognition of differences that exist; (2) increasing reading speed of trainees (at least of those that have reading deficiencies); and (3) preparation of written communication, technical manuals, visual aids, and training materials at an appropriate reading level.

The first of these implications has been treated briefly above. The second implication of reading ability (or disability) is concerned with improving the ability of slow and average readers so that they can read better and faster. The average reader reads about 300 words per minute of material of average difficulty (Buswell, 1937). After 20 hours of corrective training in reading, 62 adults increased their average reading speed from 197 words per minute to 315 words per minute, a gain of 60 per cent (Lewis, 1944). Seven out of ten slower than average readers *can* better than double their reading speed! There is much evidence back of this statement. Certain companies, notably General Electric's Lab Division at Schenectady, has a planned program to increase reading ability of employees, and other companies, particularly where job or trainee value depends partly on reading, could profit by including training in reading in their program development.

The third implication has to do with choice of word level, length, and construction of sentences in training materials. There has been a considerable amount of research in this area. Formulae have been developed for measuring the level of difficulty of training material to bring about maximum economy in training.

CONCLUSION

Worker-management relationships are considered all-important in an industrial society. Procedures for workable communications throughout the hierarchies of this society are crucial. Training is profitable to management and yields benefit to workers insofar as it engenders human understanding. When human understanding is achieved the more direct goals of management-worker morale, elimination of unauthorized work stoppages and slowdowns, higher production—are natural outcomes. Planned training may achieve these outcomes. As shown in the above discussion, personnel research and psychology have an important contribution to this purpose, worthy of management's close consideration.

READING REFERENCES

Craig, D. R., and W. W. Charters, *Personal Leadership in Industry* (2nd ed.). New York: McGraw-Hill Book Company, 1941.

Moore, B. V.; H. C. Taylor; and E. N. Hay, *Values of Psychology in Industrial Management.* New York: American Management Association, 1940.

Reitell, C., *Training Workers and Supervisors.* New York: The Ronald Press Company, 1941.

Riegel, J. W., *The Selection and Development of Prospective Foremen.* Ann Arbor: University of Michigan Press, 1941.

Starch, D., *How to Develop Your Executive Ability.* New York: Harper & Brothers, 1943.

36. FEDERAL CIVILIAN PERSONNEL ADMINISTRATION

L. J. O'ROURKE, PH.D.

Director, Psychological Institute; Chairman, Board of Advisers, Civics Research Institute. Formerly: Director of Research in Personnel Administration, US Civil Service Commission; Associate, Institute of Educational Research, Teachers College, Columbia University; Chief, Selection of Air Service Enlisted Personnel; Member, National Research Council's Committee on Classification of Military Personnel Advisory to the Adjutant General.

THE POST OFFICE STUDY

The first permanent position in applied psychology in personnel administration in the federal service, that of Director of Research of the Civil Service Commission, was set up in 1922. The first large-scale application of psychological methods to the solution of personnel administration problems was a study of the United States Post Office, which at that time numbered about 300,000 employees. This study is described at some length in order to give an insight into the procedures, the difficulties, and the reasons why results were obtained.

The first step was to secure the active cooperation of top management in the postal service, of the postal unions, and of the Civil Service Commission in a program designed to improve working conditions, efficiency, and selection methods. Such cooperation was a *sine qua non* for conducting the study.

SECURING THE CRITERION

Duties and working conditions of clerk-carriers, distributors, and railway mail clerks were studied as a basis for the criterion, for test construction, and for any possible changes in mechanical arrangements or methods that would relieve strain, improve efficiency, and result in better relations between employees' unions and management. Obviously if any changes were made in the mechanical arrangements that would affect performance of duties, they would affect selection methods. A decision on such changes would have to be made before new tests were developed, and the criterion would have to be based upon changes in procedures.

STUDY OF WORKING CONDITIONS

The preliminary study involved in part:

1. The kind and degree of proficiency necessary to perform the duties satisfactorily.

2. The qualifications necessary to attain such proficiency.

3. The qualifications necessary to learn and to qualify for the job during the training or probationary period.

4. The factors affecting relative degree of success and failure to qualify in the learning period and to acquire and maintain proficiency on the job.

5. The extent to which it would be possible to secure a valid objective criterion.

6. Insofar as a subjective criterion was necessary, the factors to be rated, how to be rated, and the extent to which supervisors would be able and willing to make valid ratings.

The analysis of duties for the purpose of developing special aptitude tests revealed methods whereby increased efficiency could be secured with less expenditure of energy. For example, an examination of the methods of distributing mail in the larger post offices showed that the distributors were throwing mail into cases arranged in front of and to the left of the workers. Since a wrist movement to the right would be easier and simpler than one to the left, it was recommended that the boxes be arranged so that mail could be thrown to the right.

Another example of ineffective arrangement was that which involved distributing second-class mail. These consisted of tiers of compartments ranged on three sides, from 6 to 10 feet high. Four men stood in front of each set and threw bundles of newspapers, periodicals, etc., into the compartments. Accuracy could be developed only after months of practice, and throwing into the high boxes was a great strain on the arms and necks of the distributors.

The Post Office had requested that tests be devised to select men that could do this work capably. But before such tests were undertaken, an analysis of the job was made, and it was suggested that a change be made in the procedure. The distributors were placed on a platform the same height as the boxes into which they threw the mail. A time study was made of the amount of mail thrown and of the number of errors made under both conditions. Under the new plan two persons could, with less strain and less expenditure of energy, accomplish as much as four had previously been able to do. Thus a task requiring considerable skill and two or three months of training was changed to a much simpler one requiring no unusual skill or endurance.

Many of the older distributors had developed varicose veins from long hours of standing. This fact caused much worry among those who felt that they would have to give up their jobs, and apprehension among others who feared that they would become similarly affected. A "leaning bench" against which the men could rest while distributing mail relieved both mental and physical strain.

Other improvements in physical layouts were made in this first study. Even today there are great possibilities for improvements in industry along these lines, that is, greater attention to the human factor in layout and design of equipment, although much progress has been made. In the federal government, particularly in the military services, psychologists have been giving considerable attention to this factor.

In the process of securing a criterion to evaluate tests for postal workers, preliminary studies were made in the post offices in Washington, D.C., and New York City. It was noted that the method of measuring output of one large group in New York really favored the careless workers, although accuracy was very important in this task. A revision of the method of evaluating output was essential in the interest of morale, for the improvement of promotion and separation procedures, and for improving the production by directing the attention of the employees to the kinds of mistakes made. Clearly a revision was essential for securing a valid criterion.

SELECTING THE EXPERIMENTAL GROUPS

Fortunately for criterion purposes, the Chicago Post Office had previously installed a system of exact timing of distributors so that the number of letters each had distributed and the number of errors each had made over a 6-month period were known. These were confidential records not known to the supervisors. The records were made available for criterion purposes. It was decided to use three separate groups of one hundred men typical of distributors in ability; i.e., each group ranging from the poorest to the best in the post office. To supplement the objective ratings, independent ratings were secured for each group from three supervisors on a form especially devised for the purpose. For each of the one hundred men selected for each of the three groups there was consistent agreement between objective ratings and supervisors' ratings.

TESTING THE TESTS

A battery of newly constructed aptitude tests was then administered to group 1. Among those included in the battery were tests designed to measure relative ability to form mental associations and memorize distribution schemes, ability to keep up scheme books, and ability to distribute mail rapidly and accurately. The type of test that had been used in the past, which consisted of letter writing, history, geography, spelling, and arithmetic items, etc., was also administered to this group. As a result of trial and statistical analysis, certain tests in the new battery were retained and certain revisions were made. The revised battery was then given to group 2. Later the same battery was given to group 3 as a check on the weightings assigned to the tests. Whereas the old test showed a differentiation a little better than chance, the new tests had a correlation of 0.71 with the criterion (O'Rourke, 1926).

SLOW-BUT-SURE FALLACY

A number of generally accepted fallacies delayed the securing of valid supervisors' ratings for criterion purposes. These fallacies were equally troublesome in securing acceptance of selection tests. The slow-but-sure concept which prevailed at that time was strikingly disproved in analysis of the test results and the actual performance of the men on monthly sorting tests which were used by the post office to determine efficiency ratings. In the sorting tests, it was found that the rapid 25 per cent of workers were 2¼ times as fast as the slow, and that the slow group made three times as many errors. A similar comparison of speed and accuracy of competitors on the new type of examination showed that the rapid 25 per cent worked 2¾ times as fast as the slow, and that the slow group made 2½ times as many errors.

In this connection it might be mentioned that a similar relation between speed and accuracy was found in a study of five hundred clerical workers which the Research Division of the Civil Service Commission also made. The 125 most rapid workers were found to work at a rate 5½ times that of the

125 workers in the slowest group, while the workers in the slowest group made 7 times as many errors as the rapid workers.

PREDICTIVE VALUE VERIFIED IN FOLLOW-UP STUDY

Less than 25 per cent of applicants for positions as distributors in the postal service are appointed. The predictive value of the new post office examination indicated that 93 per cent of those selected by means of it would be able to distribute mail with a higher degree of speed and accuracy than the average of those previously selected, as they would be able to distribute between 2 and 2½ times as rapidly, with less than one third as many errors.

A year after the new examination was first held, the Second Assistant Postmaster General wrote to all first-class post offices in the United States requesting specific reports on the speed and accuracy of men selected through the new tests. Reports, which were received from every postmaster except three who wrote that they had not completed their analysis, showed that the new men exceeded the old in speed and accuracy. Furthermore, the postmasters reported that all were able to memorize the postal schemes, and the former heavy personnel turnover had been reduced. The previous experience of the postal service had been that a large percentage of appointees had to be dismissed because even after 6 months of service they were unable to memorize the schemes.

SAVINGS IN DOLLARS AND ENERGY

The objective of the new examination was not to force employees to work beyond their normal ability. Its purpose was to select employees with aptitude for the work—employees who, when working at normal effort, would be able to memorize the schemes, keep up their distribution books, and perform their tasks with greater rapidity and accuracy. The result of the new examination was to reduce strain and turnover. Thus it had a marked effect upon the efficiency and morale of the workers selected through it.

A comparable study with an appropriate battery of tests was made for the railway mail service, and comparable results were obtained.

Assuming that the Post Office Department has paid an average of 150 million dollars a year in salaries to clerks, distributors, and railway mail clerks over the 25 years since these examinations went into effect, and assuming that efficiency was increased only 10 per cent, in place of 93 per cent, the savings in the 25 years have amounted to 175 million dollars. This should make readily apparent the dollar value of this first applied psychology program in the federal Civil Service, quite apart from the saving in energy and strain.

The cost of construction of the new examination was infinitely less than the cost of the old type. Moreover, alternate series are comparable, insuring that the competitor will receive the same score regardless of when or where he takes the test. In the past the tests varied greatly in difficulty, and the score depended to a great extent upon the judgment of the examiner.

The cost of administration of the new examination is 40 per cent less than that of the old. The Examining Division of the Civil Service Commission reported that the saving in scoring time was approximately 45 per cent. Approximately 60,000 applicants each year took the Post Office examinations 25 years ago; in later years more than 200,000 applied each time they were announced. In 1936 nearly 100,000 in New York City alone applied for the examination for railway mail clerk within 6 days after it was announced. With such large numbers, the time saved in scoring papers and making up lists of eligibles is very considerable.

OBJECTIVE VERSUS SUBJECTIVE APPROACH TO TEST EVALUATION

A belief which 25 years ago had considerable application in industry, and which is not uncommon even today, is that foremen who are thoroughly familiar with a certain phase of work are thereby qualified to construct or choose tests that will select people qualified for that work. The superiority of objective over subjective approach to test evaluation was demonstrated in the postal study. Thirty supervisors having at least 10 years experience each spent hours taking the various tests and deciding which best measured the qualities desired.

The ranking of the tests by these employees correlated negatively (—0.42) with the statistical ranking of the tests. The ranking by the thirteen employees highest in efficiency also correlated negatively (—0.38) with the statistical ranking of the tests. In this study it was possible to determine statistically the factors that influenced the judgment of the supervisors. A very high relation (0.867) was found between rankings given by supervisors to the tests and the percentage of the tests completed by them.

A negative relation (—0.467) was also found between the actual selective value of the tests and the rankings given them by other supervisors who did not take the tests but judged them merely by their resemblance to postal work. The ranking of tests in order of resemblance to postal work showed a high positive relation (0.73) with ranking by supervisors. The combination of facts indicated that judgments were influenced by the appearance of the tests and by the difficulty which the test presented to the individual making the judgment.

Tests which appear to be practical may be valuable, or they may make only a rough distinction, that is, only a separation of the very best from the very worst, or they may fail altogether to differentiate among competitors. Careful trial of all new tests, together with analysis of results, is essential to intelligent selection or rejection of tests. Such trial and analysis make possible maximum accuracy of differentiation with the use of a minimum number of tests, and therefore at a minimum cost. Unless such procedure had been followed in connection with the clerk-carrier examination, it would have seemed desirable to include three additional tests in that examination. If they had been included, they would not only have failed to increase the value of the examination, but would have increased the cost of the examination by $33,000 each year.

TESTS OF GENERAL ADAPTABILITY

For many positions in the government service a high school diploma or its "equivalent" was arbitrarily required. It was realized that state requirements for such diplomas differed widely and that high school graduates differed greatly in ability. High school graduation was required chiefly in an effort to secure persons with a certain standard of ability to learn and to adapt themselves to a position.

The Research Division undertook in 1924 to substitute for this requirement tests which would (1) measure adaptability and would permit persons to qualify who had a high degree of intelligence and aptitude but who had not finished high school; and (2) eliminate persons who had gone to high school for four years but had only a low degree of adaptability and ability to learn. At the same time it was decided to substitute this type of test for the tests then in use for many positions and which, like the former post office tests, consisted of letter writing, arithmetic, spelling, and the like.

Three power tests were devised, overlapping 50 per cent in order to cover the entire range of ability. By administering Series A, designed to replace the high school diploma requirement, it was possible to determine the average score made by high school seniors. By administering the same test to eighth grade, first-, second-, and third-year high school students, and college freshmen it was possible to determine the percentage of those who would make higher or lower scores than the average high school senior.

A score equal to that made by the average high school senior has a definite meaning, whereas it is impossible to determine from the term "high school graduation" whether an individual is below 40 per cent of seventh grade pupils or better than 70 per cent of college seniors.

The test was scaled so that each item in the test was more difficult than the one that preceded it. In 40 minutes one hundred Civil Service Commission clerks made scores ranging from 40 to 80 points out of a possible 100. A 50 per cent increase in time resulted in an average increase of only six points.

Further research disclosed that the high school requirement was too high for certain positions and too low for others. It was a simple matter to lower the entrance requirement for some positions and to raise it for others.

General adaptability tests are used for a wide variety of examinations. Insofar as a position requires proficiency in special subjects, additional tests are given to measure that proficiency. Applicants are furnished with samples and directions for each type of item so that the nature of the examination will be familiar to them.

IMPROVING STENOGRAPHIC AND TYPING EXAMINATIONS

One of the early undertakings of the Research Division was the improvement of stenography and typewriting examinations. At that time the general complaint from the departments was not that the appointees could not take dictation or type rapidly, but that they could not "use their heads" and that their English was faulty.

A general test was devised, made up of 40 judgment items, 20 non-dictated spelling items, and 20 multiple-choice English-usage items. Improved tests for measuring proficiency in typing and in shorthand were developed. These tests, together with an English-usage test developed by the Psychological Institute for a nation-wide study in English usage, were administered to all the stenographers and typists in five of America's largest industries and to 200 government stenographers working in the same stenographic pool. It was found that the Civil Service Commission was requiring too high speed in shorthand, and this requirement was lowered. The general test was made a part of the regular examination for typists and stenographers.

The scores of the 200 girls in government employment on the general test were compared with their efficiency ratings. Of those making the highest 25 per cent on the tests, 99 per cent were rated above average in efficiency. Of those in the lowest 25 per cent, only 3.6 per cent were rated above average in efficiency. This does not mean that those in the lowest group were inefficient, because they had been selected and retained as satisfactory. To differentiate among employees who are satisfactory is more difficult than to differentiate among applicants, part of whom are very low in ability.

The scores of the industrial typists and stenographers tested ranged from 1 to 85 per cent on the general test. The coefficient of correlation between their scores and efficiency ratings was 0.73. Of the 25 per cent who made the highest scores, 90 per cent were rated as above the average of the group.

One series of the general adaptability test for stenographers and the proficiency tests for typing and shorthand was released to schools and industries throughout the country. This was done as a first step in a cooperative program of government and industry designed to enable schools to meet the standards of proficiency which employers require of certain classes of employees.

RESEARCH IN ENGINEERING EXAMINATIONS

The first objective professional examination was introduced into civil service by the Research Division. This covered electrical, civil, and mechanical engineering. Previously the examination for engineers consisted of problems requiring detailed descriptive answers which required the assistance of specialists outside the Civil Service Commission for rating. This examination took four days of each competitor's time.

A new examination, constructed by the Research Division, was made up in part of problems in the form of drawings designed to measure, not only knowledge of essential facts and principles of engineering, but also ability to apply these facts and principles to practical situations. Thus the competitor would, in a problem in structural design, indicate where the strain occurred, how great the strain was, or how a structure should be made in order to stand a certain strain with a given factor of safety. These problems, as well as those of the short-answer type which they supplemented in the engineering examination, were objective in form, lent themselves to comparable scoring, and measured more directly the knowledge required for performing actual duties. They could be administered in four hours (O'Rourke, 1930).

Letters received by the Civil Service Commission in 1925 from one hundred representative universities showed that professors advised their students against entering government service. Through a cooperative program engineering professors began to recommend the service to their better qualified students. The calibre of applicants for engineering positions greatly improved. This opened up an opportunity for similar cooperation in other fields in universities, and has contributed toward securing better personnel in government.

THE COUNCIL OF PERSONNEL ADMINISTRATION

AIMS AND ORGANIZATION OF THE COUNCIL

A long step forward in the improvement of federal personnel administration was made in 1931 with the establishment of the Council of Personnel Administration. This organization consisted of the Cabinet officers and heads of certain independent establishments. The President of the Civil Service Commission was named Chairman and the Director of Research acted as Director of the Council. After two years, the Cabinet officers and other agency heads delegated to their personnel officers full authority to act as a group and to decide policy.

The functions of the Council included the development of plans and programs for building an effective career system, for devising an adequate transfer program, for improving supervision, for bringing about a closer relationship between the Commission and the various departments in such matters as improving efficiency ratings and the selection and placement of employees, for working with industrial establishments in order to bring together the best personnel practices in government and industry, for bringing about closer cooperation with state and municipal jurisdictions in personnel administration, for assisting schools throughout the country in their endeavor to train students for federal service, and for attracting and retaining better qualified personnel for government service.

A PROGRAM FOR A CAREER SYSTEM IN GOVERNMENT

One of the programs significant from the standpoint of value to government and employee alike was the Council's career and transfer program. It is generally recognized that one of the best ways of maintaining morale is to provide opportunities for advancement on merit. To have interest, effort, morale, and efficiency in government service, the service itself must be attractive. As to ability of personnel, it is clear that government cannot select any better applicants than apply. Emphasis of the Council, therefore, was directed toward attracting superior personnel and toward changing government service from the mere holding of a job to the building of a career.

In laying the foundation for an adequate career program, the Council was concerned with four essential elements:

1. A transfer system between departments and between department divisions which would make careers possible yet not interfere with the efficiency of operating units.
2. A current personnel inventory that would bring to the attention of appointing officers those employees who possessed the qualifications that would enable them to fill vacancies when they occurred.
3. The opening of definite avenues of promotion so that employees could prepare for and look forward to advancement.
4. Training and retraining for increasingly responsible positions.

PERSONNEL DATA SHEET. An integral part of the machinery for facilitating operation of the career program was the development of a personnel data sheet to contain the information necessary for making a complete inventory of the qualifications of every employee. The data were transferred to punch cards, with a master set of cards for the entire service to be kept in the Civil Service Commission. The first group of employees to be so recorded was the chemists. A trial program quickly demonstrated the practicability of filling vacancies from qualified persons already in the service.

The punch card system was the forerunner of the National Register of Scientific and Specialized Personnel, which proved of considerable value in World War II.

SAVING MONEY AND INCREASING PAY AT THE SAME TIME. A 3-month trial of the career and transfer program was held, during which only 22 per cent of vacancies in government were filled through outside appointments, 78 per cent being filled through transfer. Two outstanding examples may be cited. A call for thirty-four office workers from one agency was filled within 8 hours after the request was received. Thirty-three of these were retained permanently, there being only one replacement. A call for men with experience in both personnel work and accounting was received from another agency, with a statement that these employees were needed immediately. Candidates for positions as personnel officers and accountants had been previously interviewed. From the records thus available twelve men were placed within 48 hours without additional interviewing.

The trial clearly demonstrated that at least a half million dollars annually would be saved through such a program, and that, at the same time, there would be an increase of 1½ million dollars in employees' pay checks. This estimate does not include the worth of increased efficiency on the part of government employees because of added interest, energy, and satisfaction.

The following example shows how the filling of one vacancy in a $7,000 a year position under the Council plan could result in a saving of $2,750 a year to the government: Instead of one new appointment directly to the P-7 position, six persons are transferred—one from each grade to the next higher grade, and an original appointment is made in the lowest grade. The net result is that seven employees are receiving entrance salaries of their grades instead of salaries that, on the average, are higher than entrance salaries. The difference between average salaries of employees and entrance salaries, for the series of grades P-1 to P-7, would total approximately $2,750. Instead of saving $500—the difference between one average salary and the corresponding entrance salary—$2,750 is saved—the difference between seven average salaries and their corresponding entrance salaries; at the same time increases might be received by employees in the form of promotions and changes of grade, amounting to $2,250, and a new appointment made in the $2,000 vacancy.

The trial programs referred to were carried out at a time when federal government employees numbered less than 500,000. Today, with from three to four times as many employees, the possibilities for building a really effective career system in government are vastly greater. Personnel officers, in government as well as industry, have in general adopted the idea of the career service—of filling vacancies by promotion—but a smoothly functioning coordinated program for careers in government is still sorely needed. The groundwork was laid more than 17 years ago, and machinery for facilitating operation of a career program was devised.

PUBLIC OPINION POLL: ATTITUDE TOWARD GOVERNMENT SERVICE. In an attempt to learn specifically what factors deter the superior university student from entering government service, the Council conducted a public-opinion survey among university students in 1931. Each student was asked to list three major reasons why he would or would not be interested in a position in the federal service. The principal reasons, as shown in this survey, were as follows:

Why Interested		*Why Not Interested*	
Security of position	70%	Lack of promotion prospects	69%
Salary	53	Political influence	40
Working conditions	45	Lack of opportunity and incentive	40

It is clear that efforts to develop career possibilities in government service would tend to remove the college-trained person's chief objections to federal service (*University Training for the National Service,* University of Minnesota Press, 1932).

EFFICIENCY RATING AND SUPERVISION PROGRAMS

Two other activities of the Council of Personnel Administration that are especially significant from the psychological point of view were programs planned to aid in setting up new efficiency rating plans and to improve supervision. There was wide dissatisfaction with the efficiency rating plan formerly in use in the federal government. Industrial officials, as well as government officials, have pointed out that from 80 to 90 per cent of personnel problems can be traced to improper supervision. Useless expenditure of nervous energy, resulting from poor supervision, is a detriment to the individual and to the service.

Both the efficiency rating and the supervision programs were undertaken as cooperative efforts of the Council and the agencies. Emphasis was placed on the importance of the supervisor's assisting the employee to discover his weaknesses and improve his strong points. Good supervision should be directed, not to finding fault, but to the building of an atmosphere in which employees voluntarily point out their problems and difficulties and in which employees and supervisors jointly work out solutions.

As many as four hundred supervisors cooperated in the efficiency rating and supervision programs. Many specific weaknesses in supervision were discovered by the supervisors themselves. Wherever this program was introduced, it had much to do with raising the morale and efficiency of the employees and consequently the efficiency of the service itself.

Supervisors assisted in the criticism and analysis of the new efficiency rating forms, with the result that they were thoroughly trained in the use of the forms and understood their purposes, hence found them acceptable to a degree that could not have been hoped for under other conditions. The fact that the new efficiency rating plan was unanimously recommended for continuance without change after a year's operation was a corroboration of the wisdom of cooperative development.

The executive order establishing the original Council was revoked by President Franklin D. Roosevelt in 1938. Another organization, consisting of personnel officers and other members to be named by the President, was then set up under the same name with substantially the same objectives.

PSYCHOLOGISTS IN FEDERAL GOVERNMENT SERVICE TODAY

During World War II the number of psychologists employed by the federal government was increased many fold. In the War Department the Personnel Research Section of the Adjutant General's Office had more than one hundred and sixty psychologists working on programs that are described elsewhere in this book. A large number of psychologists worked under the Air Surgeon. The primary concern of this group was with problems of selection of potential pilots, bombardiers, navigators, and flight engineers, but its members also worked on design of controls, instrument panels, and many other problems. The Navy Department, too, employed psychologists, not only for test construction and training and personnel work but for designing equipment in relation to the human factor. The studies in the military and naval establishments involved measurement of human capacities, including resistance to heat and pressure. Some of these studies are being continued by the Army and the Navy.

Present application of psychology in the War and Navy Departments also includes the development of effective testing and evaluation programs for civilian personnel and the building of career ladders for officers and enlisted men based on new job analyses of military specialties and implemented by newly constructed proficiency examinations. Psychologists are also employed by the Services to screen new personnel to eliminate neurotics and to examine those leaving to see whether they require treatment.

The Veterans Administration employs applied psychologists for vocational counseling and guidance and clinical psychologists for work in hospitals. The work in these activities has increased to such an extent that approximately a thousand psychologists are employed by this agency. The Veterans Administration sponsors in cooperation with universities a 4-year graduate training program leading to the doctorate in clinical psychology, and pays accepted applicants salaries sufficient for self-support while in training.

READING REFERENCES

Bingham, W. V., *Aptitudes and Aptitude Testing*. New York: Harper & Brothers, 1937.

O'Rourke, L. J., *Opportunities in Government Employment* (rev. ed.). Garden City, N.Y.: Garden City Publishing Co., Inc. 1941.

———, "Saving Dollars and Energy by Personnel Research." *Journal of Personnel Research,* 1926, **9-10,** 351–364; **11,** 433–450.

Shartle, C. L., "Occupations in Psychology." *American Psychologist,* 1946, **1,** 559–582.

US Civil Service Commission, Research Division, *Annual Reports, 1923–35*. Washington, D.C.: Government Printing Office, 1935.

37. MILITARY PERSONNEL MANAGEMENT

E. DONALD SISSON, PH.D.

Research Psychologist and Assistant Chief, Personnel Research Section, AGO, Department of the Army; Member, Editorial Boards of *Personnel Psychology* and *Psychological Monographs, General and Applied;* Captain, AGD Reserve. Formerly: Director, Testing and Guidance Bureau, Louisiana State University.

The term personnel management usually refers to a multitude of activities concerned with the people of an institution or organization. Educational personnel management, for example, involves selection, counseling, evaluation, promotion, health, financial aid, vocational guidance, job placement, and a number of other student problems. Industrial personnel management, likewise, involves most of these problems and in addition such special ones as occupational analysis, job classification and placement, employee training, advancement, wage scales, morale, labor relations, and so forth. The application of scientific methods and research techniques in this important area is relatively recent—innovations made necessary by the vast organic growth of the institutions, by technological advances, and by increased specialization in curricula and jobs.

Military personnel management deals with the people who make up the fighting forces. It is concerned with most of the problems that are found in educational or industrial organizations—problems such as selection, induction, job analysis, classification, assignment, promotion, separation, morale, welfare, and safety. In the military domain, the application of psychological techniques to the solution of these problems is likewise recent, necessitated by the same factors of increased size, technological development, and job specialization. In periods of peace, these problems are pressing enough. But in war, when armies and navies are mobilizing, training, and fighting, the usual contributing factors are vastly aggravated; and to these are added the terrific urgency of speed and the vital consequences of error.

Modern warfare has been characterized by widescale mechanization and revolutionary developments in the arsenals of destruction. To stress the importance of research on armaments and matériel seems gross understatement in view of the guided missile, the proximity fuse, or the atomic bomb. But these same developments have pointed up the increased importance of human resources, and have underlined a host of problems in the realm of manpower utilization—problems to which the old-fashioned, mass-army personnel methods are as unsuited as Napoleonic strategy to a global war.

The problems of military personnel management cover a wide scope, and require for their solution the services of specialists in many fields. Occupational analysts are needed to study and describe all military jobs and to compute the services' needs in each of these. To screen and sort the incoming manpower stream requires psychometricians, interviewers, and classification specialists. To develop all of these tests, screening devices, and classification proceedings, in turn, requires test technicians and research psychologists. After classification and placement, experts in education, training aids, and methods take over. Other specialists are needed to evaluate training and proficiency. Then there are clinical psychologists for counseling services and hospitals, social psychologists to carry out opinion surveys, experts in human engineering to aid in the design of instruments and matériel, and so on through a long list.

Perhaps the largest group of military personnel psychologists are those engaged in research on the problems of selection, classification, and evaluation, and in the development of psychological instruments and procedures to implement solutions to these problems. The treatment of military personnel management which follows will be largely focussed upon the outcomes of this particular phase of the whole picture.

HISTORICAL

Modern personnel management in the armed forces had its beginnings in World War I. In the early days of that conflict, the enlisted soldier or sailor, or the draftee, was sent directly from the recruiting station or draft board to a unit needing fillers. All too often he was not adapted to this assignment. Even if he were, he might, at any time, be transferred to another outfit which had a prior claim on personnel; hence he might be taken out of the job, or the billet, in which his training had been partly completed and assigned to another unit and, in all probability, another job. Soldiers and sailors were requisitioned and assigned in mass. Placement was on the basis of the number of men needed, and assignment procedures largely ignored modern personnel principles.

THE ARMY

The work of psychologists during World War I is too well known to need detailed recounting here. The Group Examination Alpha (Army Alpha) developed then and administered to millions of soldiers was historically the forerunner of all psychological tests of the paper-and-pencil variety. The group performance test (Army Beta) was likewise a pioneer in the psychometric field. The vast amount of information assembled through the application of these tests furnished an invaluable basis for the proper classification and utilization of soldier personnel (Yerkes, 1921). No less important to the efficient handling of men were the trade tests, the job specifications, and the soldier qualification cards worked up at that time and described in the two-volume report on the Personnel System of the US Army (1919).

With the end of the war in 1918, most of the newer psychological procedures were lost sight of in the rush to return the Army to its peacetime status. During the ensuing years there were sporadic efforts in the direction of reintroducing certain of these procedures. Various short batteries of individual tests of the Binet type came into general use by recruiting officials to screen potential enlistees. In the middle twenties, the Signal Corps began employing the Otis test, and developed a test of its own for measuring aptitude for learning to receive Morse code. In connection with its technical schools for training groundcrew specialists, the Air Corps administered one of the newer revisions of the Army Alpha, and also constructed and used a number of tests of special aptitudes.

When it began to appear that a second world conflict was inevitable, those charged with planning the vast process of mobilization took steps to insure that personnel management was included at the outset. So well was this groundwork laid, and so rapid the progress, that at the outbreak of the war personnel procedures had been installed in the Army aimed at achieving the most satisfactory adjustment of the individual soldier and his most efficient utilization.

THE NAVY

Surprisingly enough, there is no evidence that the pioneer psychological methods developed in the Army during World War I were introduced into the Navy. This does not mean that nothing of this nature, however, was done in the Navy. Even before the outbreak of the World War I, a number of testing techniques were adopted for screening recruits for enlistment in the Navy and Marine Corps—all with definitely encouraging results. All of this work was decentralized, instigated by individual medical officers, and carried on under the auspices of the Medical Department.

By 1923, with the establishment of the Training Division in the Bureau of Navigation, the O'Rourke Classification Test had been adopted as the standard examination for the selection of men. The influence of this screening test can be traced in the gradually improved calibre of the new recruits. So successful was it, in fact, that its use at recruiting stations became regularly established practice up until World War II.

The period of the second war saw the application of scientific personnel procedures on the widest possible scale. In all of the armed forces—Army, Navy, Marines, Coast Guard, and in the aviation arms of all forces—modern methods of selection, classification, and assignment were explored to the fullest to effect the most economical utilization of manpower resources. It is impossible in the limited scope of the present section to examine all of the various military personnel programs in any detail. Nor is it feasible to pursue all of the short shifts and turns in the development of any of these programs. Attention, therefore, will be focused in the main on those programs which affected the largest numbers of individuals—the enlisted personnel programs of the Army and Navy exclusive of the aviation arms. In limiting our treatment in this fashion, it is not intended to convey the impression that these two approaches to the problem were more scientific, more highly developed, or more adequate from the point of view of proper personnel practices. Nor is it intended to imply that in these two systems the actual fulfillment always realized the complete expectations of management policy and theory. There were mistakes, to be sure. Sometimes these were colossal, and served as dramatic monuments of unforgivable and profligate waste of our human resources. More often, however, there were scattered but striking instances of seemingly poor classification, such as civilian truck drivers assigned as cooks, or men with civilian experience as cooks assigned to learn the duties of radio operators. Important as they were to the individuals concerned, these examples were numerically insignificant against the background of millions of soldiers and sailors. And yet, wherever such instances occurred because of faithful adherence to classification regulations, rather than in spite of these regulations, as was often the case, they represent an important but widely overlooked characteristic of wartime personnel procedures: that is, that men had of necessity to be classified and assigned not always to the job for which they were best qualified by skills, aptitudes, interests, and experiences, but rather to a job which needed doing, and which they could reasonably be expected to do. When the Army needed radio operators, for example, men who showed by test performance an aptitude for this kind of work had to be selected and assigned, whether or not they had had previous related experience, and whether or not they might serve better in some other job for which the immediate need was less crucial. The fact that selection and classification were decentralized to the various centers where training was carried out, and that quotas for needed specialists were allotted and suballotted to these centers and units, only served to aggravate the situation. In a word, wartime selection and classification in the armed forces were akin to vocational selection rather than vocational guidance.

ARMY PERSONNEL PROCEDURES

SELECTION

Not every man who was reported to the various Armed Forces Induction Stations by the Local Boards of the Selective Service System was proper material for the military service. Some were too old or too young; some were unsuited by reason of poor moral character; others had medical defects (including psychoneurotic or psychotic complications) which disqualified them for the rigors of Army life. Finally, there were those whose educational and intellectual deficiencies would have necessitated an inordinate expenditure of training time and training personnel. In addition to administrative and medical examining, therefore, it was necessary to insure that all men finally accepted for service in the armed forces possessed the minimum mental standards for induction.

These minimum standards were defined in terms of successful performance on a battery of objective psychological tests which were administered to all nonhigh-school graduates. The purpose of these tests was first to screen out the illiterate or marginally literate (and the non-English-speaking) selectees, and second, to identify in this group those with sufficient potential ability to justify their induction. The first hurdle in the induction station testing battery was the Qualification Test—a short test designed to screen out those men whose intelligence or literacy level, or both, were insufficient to recommend immediate assignment to regular basic training.

The test required only about fifteen minutes to administer. It was composed of seventeen questions of the free-answer type dealing with simple arithmetic, understanding of directions, paragraph reading, and the like.

Those who failed the Qualification Test were given next a group aptitude test called the Group Target Test, which involved perception and memory of patterns, pattern orientation, and directional orientation. This test required no reading or writing. The examiner indicated certain prescribed patterns or movements on a large wall-chart, together with simple verbal instructions. The examinee responded by drawing lines connecting the proper dots on a small replica of the wall chart. This test was chosen because it was found to correlate well with ratings of "soldier performance" in a training situation.

Those who passed the Group Target Test were recommended for induction. Those who failed were given a final Individual Examination involving timing and coordination, block patterns, reproduction of design, and simple directions. Failure on this final hurdle was considered sufficient basis for rejection. Those who passed, along with those who performed successfully on the Group Target Test, were sent to Special Training Units for an intensive period of instruction in the fundamentals of reading and arithmetic. After this, if a satisfactory level of literacy had been achieved, they were assigned to the regular training regime.

CLASSIFICATION

From the induction station, the new soldier went next to a reception center where he was given a uniform, was innoculated, oriented, tested, interviewed and, in general, afforded his first real taste of Army life.

THE CLASSIFICATION INTERVIEW. Each man was interviewed individually by a trained Personnel Technician for the purpose of discovering and recording all information about him that should be considered in classifying and assigning him to his proper job in the Army. The various items of information obtained through this interview included age, physical condition, education, civilian occupational experience, foreign language skills, hobbies, previous military experience, and so forth. The phase dealing with occupational history was especially important, since it was essential to know precisely what each man had done and to spot certain critical occupational skills whenever they turned up. In this connection, the interviewer made full use of the job descriptions contained in the "Civilian Occupational Classification of Enlisted Personnel" (War Department Technical Manual 12-426, 1944) supplemented, for the less common civilian jobs, by the "Dictionary of Occupational Titles." To check on the soldier's claimed occupational skill—to discover whether he was as well informed as the average expert in the particular field or whether he possessed only very little relevant information—the interviewer had access to the "Oral Trade Questions" developed by the United States Employment Service.

During the interview, all important information elicited was entered on a special Soldier's Qualification Card (WD AGO Form 20). Afterwards, the entries for primary and secondary civilian occupation were classified and an estimate of the soldier's degree of skill in these occupations was arrived at. Finally, most of the information on the card was coded and punched—in specified sections along the edges—in such manner as to permit convenient sorting according to any desired variable.

The Form 20 initiated during the initial classification interview remained with the soldier's unit throughout his entire Army career, and became a cumulative record of all of his service, his test scores, qualifications, and experiences.

CLASSIFICATION TESTS

Among the items of information recorded on the Form 20 were scores obtained on all tests administered to the soldier. The first of these tests which he encountered was the Army General Classification Test, administered during the war to nearly 10 million men. The AGCT is a conventional test of "general learning ability" or intelligence. It was produced in various parallel and alternate forms—AGCT-1a, 1b, 1c, and 1d produced in 1940 and 1941, and AGCT-3a and 3b produced in 1945 and 1946. Each of the four forms of AGCT-1 contained three types of items in spiral-omnibus arrangement: vocabulary, arithmetic, and block-counting (see examples on following page). AGCT-3 consists of four subtests separately administered and separately scored: Reading and Vocabulary, Arithmetic Computation, Arithmetic Reasoning, and Pattern Analysis. The content of the first three subtests is clear from the titles; the last (Pattern Analysis) consisted of sets of line-drawn patterns and forms and was designed to measure skill in the mental manipulation of spatial relations and the visualization of three-dimensional form. The total of the four sub-

Samples of Test Items from the Army General Classification Test

A. Vocabulary Items

43. A CRAFTY scheme is one which is
(A) meager (B) difficult (C) sly (D) artistic

44. A company commander RECOMMENDS a man when he
(A) commands him (B) praises him (C) corrects him (D) reproaches him

45. A DETOUR is a
(A) man at work (B) bridge out (C) road machine (D) by-pass

B. Arithmetic Items

33. To find how many are 8 yards of rope and 9 yards of rope would you
(A) add (B) subtract (C) multiply (D) divide

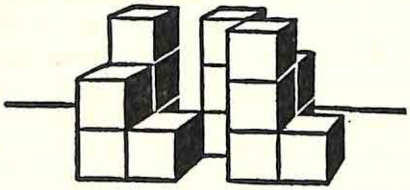
34. Eight men went on a trip by automobile. The total expense was $18.00 which was shared equally. To find each man's share would you
(A) add (B) subtract (C) multiply (D) divide

35. A train was going to Cleveland which was 85 miles away. It broke down 20 miles from Cleveland. To find how far it had traveled would you
(A) add (B) subtract (C) multiply (D) divide

36. A camp has 316 men in 4 equal groups. To find how many men are in each group would you
(A) add (B) subtract (C) multiply (D) divide

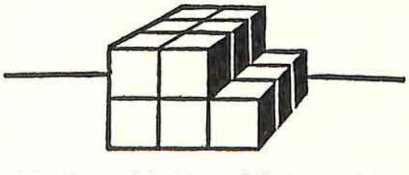
C. Block Items

37. How many boxes?



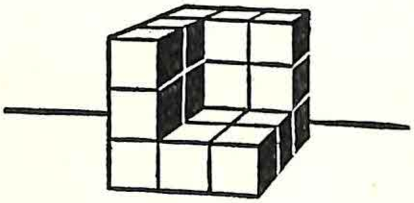
(A) 17 (B) 15 (C) 16 (D) 14

38. How many boxes?



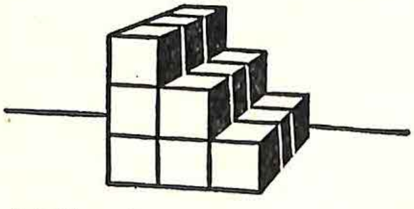
(A) 14 (B) 11 (C) 15 (D) 16

40. How many boxes?



(A) 20 (B) 19 (C) 21 (D) 15

41. How many boxes?



(A) 15 (B) 9 (C) 18 (D) 12

tests yields an over-all score equivalent to those obtained with the four earlier forms of the test.

In order to make scores on each of the forms of the AGCT comparable, and to standardize the recording and interpretation of test performance throughout the Army, raw scores were converted into Army Standard Scores. The Army Standard Score scale had a mean of one hundred and a standard deviation of twenty, based on the total population of selectees. Actually, the standardization was carried out, and the standard score scale constructed, before the selectees began to arrive at Army reception centers. It was, therefore, necessary to estimate the characteristics of the population of potential draftees. This estimate was not perfect, and the subsequent distribution of standard scores that resulted as the numbers of soldiers tested grew into the millions was somewhat skewed.

For many purposes it was more convenient, and somewhat simpler, to report AGCT scores in Army grades. These five grades were defined in terms of the standard deviation of the distribution of scores and were originally symmetrical with respect to the mean. Because of the skewness of the obtained distribution, however, the lower limit of Grade IV was extended downward an additional half standard deviation. The actual grade distribution of all men processed through reception centers through May, 1946, is shown in Table 1.

Table 1: AGCT Grade Distribution of Men Tested Through Reception Centers, 1940–46

$N = 9{,}757{,}583$

Army Grade	Standard Score Limits	Percentage
I	130 and above	5.8%
II	110 — 129	26.0
III	90 — 109	30.6
IV	60 — 89	28.8
V	59 and below	8.8

It has been stated that the AGCT was a test of "general learning ability." This was the terminology used in official documents (War Department *Technical Manual* 12-260, 1946). The existence of a general trait of this nature may be questioned; in any event, the AGCT measures much the same thing as other well-known tests of mental ability or intelligence. This is illustrated in the correlations presented in Table 2 (Staff, Personnel Research Section, AGO, 1945).

Table 2: Correlations between AGCT and Other Tests of Mental Ability

Tests	*N*	AGCT *Means*	AGCT *SD*	*r*
Army Alpha (Wells' Revision—long form)	750	98.3	22.7	.90
Otis Higher Mental Ability Examination	1646	103.5	23.1	.83
ACE Psychological Examination	1371	114.1	16.3	.79

Whatever it measures, whether "general learning ability," mental ability or what-not, the AGCT was of value to the Army's personnel system because it predicted fairly well how soldiers would perform in a wide variety of training courses and job assignments. The various Army jobs (Military Occupational Specialties) varied over a considerable range in the extent of training they required, and this was reflected in the average AGCT Scores of men qualified in these jobs. This relationship between MOS and average test score is shown, for several selected jobs, in Table 3 (Sisson, 1947). In this Table, the column *SSN* indicates the Specification Serial Number of the MOS.

Table 3: Average AGCT Scores for Various Military Occupational Specialties

SSN	*MOS*	*Average AGCT Score*
787	Weather Forecaster	134
275	Classification Specialist	124
647	Radio Repairman—Aircraft Equipment	120
667	Message Center Clerk	115
660	Tank Mechanic, Minor Maintenance	105
566	Duty NCO	101
824	Mess Sergeant	100
677	Military Policeman	98
238	Lineman, Telephone and Telegraph	92
060	Cook	87
188	Duty Soldier II	80
271	Longshoreman	72

It must be remembered, of course, that these figures are averages, and do not indicate the wide variability of AGCT scores in any one occupational category. The lowest tenth of men qualified as "Classification Specialists," for example, scored below a standard score of 109, whereas the top tenth of cooks scored above 114. Nevertheless, the AGCT was a valid predictor of achievement in training courses of many sorts, as is shown in Table 4 (Staff, Personnel Research Section, AGO, 1945). It found widespread use as a requirement for most specialist courses in the Army. A score of 100, for example, was required by many service schools; applicants for Officer Candidate training needed scores of 110 or above; and trainees of the Army specialized Training Program had to achieve a qualifying score of 115.

Table 4: Validity of the AGCT for Predicting Grades in Training Courses

Course	*N*	*Mean*	*SD*	*r*
Administrative Clerical Training, (AAF)	2947	121.7	11.1	.40
Clerical Training, WAAC	199	116.8	12.0	.62
Airplane Mechanics	3081	118.1	10.7	.35
Motor Mechanics	318	88.3	24.4	.69
Radio Mechanics, AAF	419	108.0	13.0	.49
Bombardier Training, AAF	40	111.5	18.6	.62
Intelligence Training, AAF	104	118.9	10.6	.51
Weather Observers, AAF	1042	130.2	12.5	.43

The second test administered at the reception center was the *General Mechanical Aptitude Test.* This test was composed of three parts: mechanical information, surface development (similar to the pattern analysis section of the AGCT-3), and mechanical comprehension. Each part was administered with a time-limit of fifteen minutes; the composite score was used to predict success in mechanical training courses and mechanical job assignments. Some idea of the validity of the test for this purpose is indicated in Table 5, which is adapted from War Department *Technical Manual* 12-260 (1946).

Table 5: Validity of the General Mechanical Aptitude Test for Various Training Courses

Course	*N*	*Mean*	*SD*	*r*
Automotive Mechanics	330	105.0	17.2	.54
Airplane & Engine Mechanics	453	106.3	15.0	.45
Airplane Armorer	102	106.8	14.5	.51

The final test administered to the soldier at the reception center was the *Army Radio Code Aptitude Test.* Modern armies are seriously dependent upon radio communications for the control and coordination of their far-flung operations. The Army required a far greater number of radio operators than it received through the draft, and consequently the training of these operators was an urgent necessity. Experience with this training brought out vividly the well-known facts of individual differences in skill and aptitude. Table 6 shows the variability in performance for one class of radio code operator trainees after eight and twelve weeks of training (adapted from T.M. 12-260, 1946).

Table 6: Code Receiving Performance of Trainees after 8 and 12 Weeks of Training

Words Received per Minute	Number of Trainees After 8 Weeks	After 12 Weeks
24–26		1
21–23		0
18–20	5	8
15–17	1	8
12–14	20	27
9–11	11	19
6–8	31	17
3–5	19	6
0–2	8	1
Dropped		8
Total	95	95

In view of the wide variability evidenced here, it was clearly essential to select for training, if possible, those whose expectations of satisfactory progress were greatest. The Army Radio Code Aptitude Test was used for this purpose. By means of phonographic transcriptions, it first offered paired presentations of code signals and letter identifications, for a number of learning trials, and then, in the test itself, presented the code signals alone for the examinee to identify. The validity of the test is shown in the following figures which indicate the percentage of persons at various score ranges who were able to receive code signals at the rate of twelve words per minute or better after two months of training (War Department *Technical Manual* 12-260, 1946).

Army Standard Score ARC-1	*Per Cent Receiving at a Rate of 12 WPM or Better*
130 and above	87
110 — 129	70
100 — 109	48
Below 100	32

From the reception center, the soldier was assigned to one of the training centers where he underwent basic military and technical training. The process of classification, of course, continued. When, at the end of the training cycle, it became necessary to select men for advanced technical training in one of the service schools, or to determine the job classification and assignment of the soldier, the information recorded on the Form 20, including the scores on the tests given at the reception center, was consulted. There were also available a number of additional tests of special aptitudes or job proficiency, including the following which are largely self-explanatory:

Aircraft Warning Tests
Army Individual Test
Auto Mechanics Test
Carpenter Test
Clerical Aptitude Test
Clerical Achievement Test
Cooking Test
Cryptography Test
Dictation Test
Distributor and Valves Test
General Technical Test
Machinist Test
Nut and Bolt Manual Dexterity Test
Supply Clerk Test
Technical Trade Test
Tool Usage Film Strip Test
Trade Information Test
Truck Driver Test
Typing Test
U-Bolt Assembly Test
Use-of-Tools Test
Weather Aptitude Test
Welding Test

OFFICER CANDIDATES

The need for commissioned officers was pressing during World War II, and officer-candidate schools were established in each of the arms and services to fill this need. These schools offered formal training in the duties and responsibilities of officers, and provided for a period of intensive observation and screening of the candidate-trainees. Failures in training were about equally divided between academic failures and those due to deficiencies in "leadership" qualities. To reduce the first category of failures, insofar as possible, it was required that each applicant have a score of 110 or above on the AGCT and, later, a score of 115 or above on the Officer Candidate Test. This test contains items of paragraph reading, interpretation of graphic material, arithmetic reasoning, and other types, with item difficulties somewhat greater than those of the AGCT. In several studies of officer candidate populations, the correlation between test score and academic grades varied around .60 and .70—a validity of considerable size, especially when it is remembered that the populations involved were rather severely restricted in range of the abilities measured (mean standard score —125.2; standard deviation —12.4) through prior selection by means of the AGCT.

The evaluation of leadership and personality characteristics was attempted during the war by means of the oral examination or interview conducted by a board of officers. Evidence concerning the validity of this traditional method of evaluating fitness to become an officer was largely negative, and the procedure was replaced, after the end of hostilities, by better predictive instruments including a scorable self-description blank, an efficiency rating scale, and a more objective standard-type interview.

OFFICER SELECTION

Toward the close of World War II, the Army conducted a research program for the development of tools and procedures for the selection of several thousand wartime temporary officers for retention in the Regular Army. This research covered all phases of the problem and resulted in the development of several new personnel selection devices and procedures including:

The General Survey Test for determination of the general level of educational achievement
The Biographical Information Blank—a scorable self-description form
The Officer Evaluation Report—a rating form embodying several types of scales including the "forced-choice" rating scale
The Structured Committee Interview—a standardized interviewing procedure employing work sheets and scorable report forms

Two of these devices have been described in publications: The Forced-Choice rating procedure by Sisson (1948) and Bittner (1949) and The Structured Interview by Rundquist (1947).

In order to validate this battery, a criterion of officer effectiveness was developed. More than some 15,000 officers, meeting in small groups of closely associated co-workers, rated each other by "nominating" their most competent and least competent fellow officers.

Where there was substantially unanimous agreement that officers were outstandingly high, average, or outstandingly low in competence, these officers were placed in high, middle, and low criterion groups. The battery of testing instruments listed above was administered to these groups and their scores compared with their criterion standings.

With scores in the last three instruments—Officer Evaluation Report, Biographical Information Blank, and Interview—combined, the results shown in Table 7 were obtained (Richardson, 1947).

Table 7: Validity of Battery Composite score for Selecting Officers

Selected on Composite Score Per Cent		Percentage in Each Criterion Group Low	Middle	High
Top	30	2	24	74
Middle	40	24	44	32
Bottom	30	71	21	8

NAVY PERSONNEL PROCEDURES

SELECTION

Prior to, and during the early stages of the war, enlisted personnel for the Navy were procured through voluntary enlistments, and recruits were required to satisfy certain requirements on the Navy General Classification Test and to meet a specified educational requirement. After Pearl Harbor, personnel were procured largely through Selective Service, and through the operation of the Armed Forces Induction Stations. Personnel for all of the armed forces—Army, Navy, Marines, and Coast Guard—were screened by means of the same instruments described above—Qualification Test, Group Target Test, and Individual Examination. The inductee's preference for one of the services was considered when possible; otherwise men were allotted by quota. At the same time, naval recruiting stations continued to enlist volunteers below minimum draft age (17 year olds) for special programs, and for this purpose employed other selection devices such as the Radio-Technician Selection Test.

CLASSIFICATION

From induction stations and recruiting stations new men were sent to recruit-training commands or "boot camps" where they were outfitted, oriented, and classified, and where they were given their basic or "boot" training. Classification here served the same purposes as in the Army training center—to gather together all pertinent information about the man in order to effect his maximal usefulness to the Navy through assignment to one of the elementary naval training schools, to a special billet, or to duty as a "general detail" hand.

THE BASIC TEST BATTERY. As a first step in the classification process, all men were administered the Basic Test Battery. This group of tests was developed during the early part of the war to provide coverage of the skills and aptitudes involved in the more than 800 different types of billets to which the men might eventually be assigned. Somewhat more extensive than the initial classification battery employed by the Army during wartime, the Basic Test Battery consisted of eight different tests, involving over eight hundred test items, and required some three and one-half hours to administer.

The first of these tests in the battery was the General Classification Test. Unlike the Army GCT, this Navy test included items involving verbal functions only. These items were of three types: sentence completion, opposites, and analogies (see page 245). The test contained 100 items; raw scores were converted for reporting and recording purposes, into a standard score scale, derived from the total distribution of Navy enlisted men, and having a mean score of 50 and a standard deviation of 10.

As might be expected, the GCT proved to be a valid predictor of course grades in a wide variety of training courses as shown in Table 8 (Staff, Test and Research Section, Bureau of Naval Personnel, 1945). The validity is considerably higher for those courses involving verbal ability—radar operators, yeomen (clerks) etc.—than for those largely me-

Table 8: Estimated Coefficients of Correlation between Navy GCT Scores and Final Grades in Ten Types of Elementary Naval Training Schools

Type of Training Course	*N*	*Corrected r**
Basic Engineering	1480	.52
Diesel	2160	.42
Electrical	1747	.52
Gunner's Mates	1677	.38
Radar Operators	1053	.60
Signal	984	.49
Store Keepers (WAVES)	678	.44
Torpedo	880	.32
Yeomen (WAVES)	738	.62

* *Corrected for restriction in range of test scores.*

Samples of Test Items from the Navy General Classification Test

A. Sentence Completion

C. I thought he was asleep because his eyes were . . . .
1 dark 2 shut 3 dull 4 gray 5 heavy

D. A good sailor will . . . . the orders of his superior officers.
1 see 2 fear 3 read 4 obey 5 like

E. He awoke early in the morning and saw the . . . . sinking below the horizon in the west.
1 sun 2 moon 3 storm 4 waves 5 clouds

B. Opposites

H. SELFISH
1 generous 2 lazy 3 playful 4 happy 5 beautiful

I. BREAK
1 weep 2 find 3 handle 4 use 5 mend

J. GRIEF
1 anger 2 poverty 3 joy 4 sorrow 5 pride

C. Analogies

M. FISH is to *swim* as BIRD is to
1 air 2 fly 3 feathers 4 ride 5 wings

N. WATER is to *sponge* as INK is to
1 pen 2 bottle 3 write 4 blotter 5 desk

O. THERMOMETER is to *temperature* as CLOCK is to
1 month 2 heat 3 time 4 hand 5 dial

chanical in nature—torpedo, machinist's mates, and gunner's mates.

The second test in the battery, the Reading Test, consists of six paragraphs of increasingly difficult text, each followed by a number of questions based on the paragraph. In general, this test yields validity coefficients of about the same order as the GCT, that is, high for schools requiring verbal ability and the use of symbols and low for those involving emphasis on mechanical subjects.

The Arithmetic Reasoning Test contains thirty items designed to evaluate the examinee's quantitative reasoning ability. Its validity for predicting grades in elementary schools—like the two earlier tests—is high for yeomen, radar operator, and basic engineering courses (corrected *r's* of .63, .61 and .63, respectively) and low for torpedo, gunner's mates, and diesel courses (corrected *r's* of .28, 31 and 36, respectively).

The Mechanical Aptitude Test contains three sections: Block Counting, Mechanical Comprehension and Surface Development (similar to the Pattern Analysis of the Army's AGCT-3). The verbal element is minimized in this test—most of the items are "picture" items—and the spatial factor, which is known to be an important ingredient of mechanical aptitude, is heavily represented in the first and third sections (block counting and surface development). Validities for the test, against grades in elementary training schools, were reasonably high for such courses as basic engineering, radar operation, and machinists' mates (corrected *r's* of .52, .50 and .48, respectively).

The next test in the Basic Battery, the Mechanical Knowledge Test, is essentially two tests in one, yielding two separate scores—mechanical and electrical. The first part of the test includes picture items involving the identification of relationships or matching of parts. The second section includes verbally presented information items. Both sections cover the mechanical and electrical areas. Scores on these two areas were useful in predicting success in such primary training courses as Aviation Machinist's Mate, Basic Engineering, Diesel and Electrical.

A Clerical Aptitude Test was included in the battery to cover the aptitude factors presumably important in storekeeper and yeomen billets. It consisted of three parts: Alphabetizing, Name Checking, and Number Checking. The test proved to be a useful prognosticator, not only for the two schools mentioned, but for quartermaster and signal schools as well.

A Spelling Test which required the identification of the one word in five of each item that was spelled incorrectly proved to overlap too much with the GCT, and to have a predictive value too low to justify its continued use.

Finally there is the Radio Code Test—Speed of Response. None of other tests of the Basic Battery had been of much value in predicting success in code-learning classes. The selection and training problem was critical since failures at these schools were many times higher than at other schools. It was not unusual, for example, for as many as a third of beginning radio trainees to fail. The Radio Code Test—Speed of Response was developed for the armed forces by the National Defense Research Committee. Like the Army Radio Code Aptitude Test, previously mentioned, it involved the learning of three code symbols, through phonographic presentation. These same symbols were then presented at more rapid speeds, and the examinee was required to identify as many as he could. The validity of the test was moderately high (corrected $r = .39$) against final grades in radio courses. While not as high as might be desired, this was higher than any other predictor for this course; and experience with the test (and a fairly high selection ratio) showed that attrition in radio training could be reduced by nearly a half.

In addition to taking the tests of the Basic Battery, each man was interviewed, and all pertinent information obtained in this way. All test scores were recorded on the Enlisted Personnel Qualification Card which became a part of each individual's permanent service record.

From the recruit training command, after completion of "boot" training, approximately 40 per cent of the men were sent to the various elementary training schools. Approximately 10 per cent were assigned to special billets, given "rates" or directly commissioned, on the basis of previous experience or special talent. The remaining half of the men were assigned to duty as "general detail" hands.

Many of these latter, as well as some of the graduates of elementary training courses, were assigned to precommissioning centers where they were further classified and sorted to make up the complements of new construction vessels. Classification procedures at this point usually involved "balancing" the crews of the various ships to be commissioned, drawing up billet specifications for the many jobs to be done, and organizing personnel information files so that proper assignments to these billets could be made. The continuing classification process was also focused at receiving stations, where men were assigned (as casuals or transients) for reassignment to new ships or to shore stations, and at naval districts, shore stations, service schools, and a few of the larger ships. In all of this work the information and test scores recorded in the Qualification Card were available. In addition a number of special tests were employed for selection for certain billets. These tests included:

The CIC (Combat Information Center) Aptitude Test
The Pre-Radar Officer Aptitude Test Battery
The Radio Technician Selection Test
The Sonar Pitch Memory Test
The Mechanical Comprehension Test
The Digit Memory Span and Sentences in Noise Tests for Telephone Talkers
The Airplane Matching Test
The Enlisted Personal Inventory
The Billet Qualification Blank

SUMMARY

It is clear from the account set forth in the preceding pages that scientific personnel techniques and procedures were developed and adopted on an unprecedented scale by the Army and Navy. This was also true to the same extent of the aviation arms of the two major forces and to a slightly lesser degree, of the other armed forces—the Marine Corps and the Coast Guard. It is unfortunate that it has not been possible, in the space available here, to discuss in detail the selection and classification programs of these other elements. The Aviation Psychology Program of the Army Air Forces, for example, involved wide-scale research activities concerned with the selection and classification of air crews—pilots, bombardiers, navigators, etc. This program alone involved a staff of more than one thousand people, of whom approximately half had received professional graduate training in psychology. The reports on its many research activities have filled nineteen volumes (Army Air Force, 1947).

Once before, during World War I, psychologists left their classrooms and laboratories and applied their scientific techniques to the problems of the effective utilization of personnel in the Army. The outcomes of their work were far-reaching, yet their products were largely abandoned by the services on the return to a peacetime status. It is most encouraging to note that this error has not been repeated a second time, and that psychological techniques continue to play an important part in the personnel programs of all the military forces.

READING REFERENCES

Army Air Forces, Aviation Psychology Program, Research Reports Nos. 1–19. Washington, D.C.: Superintendent of Documents, 1947.

Bingham, W. V., "Military Psychology in War and Peace." *Science,* 1947, **106,** 155–160.

Bureau of Naval Personnel, Test and Research Section, in cooperation with NDRC Project N-106 and the College Entrance Examination Board, *Personnel Research and Test Development in the Bureau of Naval Personnel.* Princeton, N.J.: Princeton University Press, 1947.

Vernon, P. E., "Research on Personnel Selection in the Royal Navy and the British Army." *American Psychologist,* 1947, **2,** 35–51.

War Department, *Civilian Occupational Classification of Enlisted Personnel.* Washington, D.C.: Government Printing Office, War Department Technical Manual No. 12–406, 1944.

38. METROPOLITAN POLICE PERSONNEL

SADIE MYERS SHELLOW, PH.D.

Consulting Psychologist in Private Practice and Psychologist for the Jewish Vocational Service, Milwaukee, Wisconsin; Lecturer, University of Wisconsin School for Workers. Formerly: Clinical Psychologist for the Juvenile Psychopathic Institute of Chicago; Industrial Psychologist for the Milwaukee Electric Railway and Light Company, and Psychologist, Milwaukee Police Department; Consulting Psychologist, Rhea Manufacturing Company.

The modern metropolitan police department is a complex organization employing large numbers of men with diverse functions. But until comparatively recent years, little or no attention was paid to the personnel problems inherent in the administration of police work. The original and often sole function of police departments was the apprehension of criminals. Many new police functions have developed which are essential to the orderly life of a community, e.g., regulation of traffic and accident prevention, prevention of crime, handling of juvenile delinquents, public relations, control of strikes and riots, investigation of subversive activities, licensing of taverns and lodging houses, and regulating domestic disputes. To cope with these many new problems, the police officer needs to be trained in improved methods and given information quite different from that required for his original function.

The use of scientific laboratories and techniques in investigation, the constant increase of clerical work, records, and files, and the use of cars and motorcycles have also greatly changed the character of police work. Today the aim of the administrators of police is to professionalize the service. Proceedings of the meetings of the International Association of Police Chiefs indicate the widespread concern for raising standards among officers and improving the attitude of the public with regard to police work. Since 90 per cent of the police budget is spent upon personnel, the problem of improving police service is to a large extent a personnel problem.

PSYCHOLOGICAL SERVICES

The actual personnel organization of a police department is similar to that in other public institutions, although casual acquaintance with it leads one to think of it as vastly different. The personnel work of a police department includes (1) selection techniques for recruits, sergeants, lieutenants, captains, instructors, and for the various types of jobs performed by patrolmen, e.g., motorcycle men, squad car men, detectives, assistants in laboratory, and clerks; (2) methods for training new men and for training men in service for promotion, for instructors, for the supervisory force, and for specialized jobs; (3) follow-up procedures for determining efficiency; (4) clinical study of recruits by interview upon selection and completion of training programs, and similarly, perhaps, at end of one year of service; (5) also a clinical study of any officer involved in an accident, of any problem referred by the commanding officer, of all men considered for promotion, and of any man who may voluntarily come for advice; (6) research along such lines as the development of psychological tests, establishing relationships between tests, initiating training programs, and preparing service records; and (7) public relations between social agencies and the police department.

SELECTION

Many police departments are handicapped by outmoded methods of recruiting, such as political appointment on recommendations from ward politicians or appointments by politically appointed police commissions. Various departments have felt the need for more objective standards and have called in psychologists to develop tests and measuring devices to aid in the selection of recruits. Owing to a lack of a complete job analysis of the functions and duties of a police officer, early investigators necessarily selected tests based upon their interpretations of the qualities required of a police officer.

MODERN TESTING

Psychologists have used a variety of tests in attempting to discover the essential qualifications for police officers. The General Adaptability Test (O'Rourke, 1933) was prepared at the request of the International Association of Chiefs of Police and has widespread use in metropolitan police departments. It measures the ability to reason, to understand and interpret directions. Its advantage over ordinary intelligence tests lies in its use of problems which pertain to police work but which do not require previous police training in order to be answered successfully.

Since police personnel is now selected in three fourths of all cities of the United States by Civil Service Commissions (Smith, 1940), the elimination of unfit officers must be done by police departments during training or in the probationary period for recruits. Individual studies are made and various tests administered during this time and the results are used as a basis of rejection or assignment to jobs. Psychological tests of the following categories are recommended for investigating the qualifications of these recruits: (1) intelligence, (2) general adaptability, (3) emotional stability, (4) social attitudes, (5) initiative, and (6) resourcefulness. Examinations made by experts other than psychologists include the following qualifications: (1) character, (2) physical condition, and (3) psychiatric condition.

SUPERVISORY PERSONNEL

Sergeants, lieutenants, captains, and other supervisory personnel have additional duties involving public relations, supervision of the regular officers, and maintaining the morale and efficiency of the department. Here the qualifications which apply in industry to a good foreman are necessary, but thus far no satisfactory selection procedures have been devised in police departments for promotion of men from the ranks to these positions.

SPECIAL SERVICES

Because of the growing specialization in modern metropolitan police departments, police officers need to be selected on the basis of the particular job to which they are assigned. The regular patrolman has duties peculiar to the patrolling of his beat; the officer who operates a squad car requires different qualifications; specialized qualifications for motorcycle officers and detectives must be considered. There are various jobs for which selection procedures are needed, such as laboratory assistants and special investigators. Selection for motor vehicle operation illustrates the adapting of psychological testing techniques to police department personnel work.

Drivers' tests are used to discover slow reaction time, distraction, distribution of attention, width of vision, and safety under emergency conditions. Not only should a motorcycle man be tested before being assigned to a motorcycle job, but examined periodically because many motorcycle officers must be taken off duty for a period to relieve tension and fatigue from vibration of the machine.

TRAINING

Prior to 1920 there was practically no systematic training for police officers in the United States. New recruits originally were broken in by older men, but better selection has been followed by a recognition of the need for improved training programs. Now, practically all recruits receive organized training, extending from thirty days in certain cities to a program in others which is continuous throughout the entire length of service. Because of the enlargement of the function of the Federal Bureau of Investigation, efforts are being made to project some measure of consistency into police training. At present there are six principal types of organization for providing training for police service: (1) the National Police Academy of the Federal Bureau of Investigation, (2) state police training programs, (3) zone schools within a state, (4) short intensive courses or institutes, (5) itinerant instructors, and (6) local police training programs. In some of these programs psychologists have participated along with other specialists, but the psychologist functions best as a regular staff officer and instructor in psychological problems in the local training school.

Success as a police officer depends not only on a knowledge of city ordinances and police regulations, but also upon an understanding of human motives and individuality of personalities. No other single group of public servants aside from teachers and clinicians encounters as many problems of human behavior so constantly and intimately as do police officers. In one typical metropolitan city of more than 500,000 approximately 6,000 domestic relations problems alone are handled during the year by the police officers and less than 5 per cent of these are taken into court. A detailed job analysis of the police officers' duties made by the US Department of Interior (1938) indicates the need for instruction of police officers in such well-known psychological fields as defects and inaccuracies of memory, memory training, illusions, errors of perception, attention span, observation, judgment, psychology of testimony, mental hygiene, public relations, and behavior problems both among children and adults. Often much of the validity of a patrolman's testimony is based upon his recognition of the limitations of sensory perception as well as his keenness and accuracy of observation.

PLANNED TRAINING PROGRAM

One of the earliest efforts to establish a planned training program was made by Police Chief Laubenheimer of Milwaukee, who set up a training school which was unique in that it was continuous. Chief Laubenheimer believed that new conditions were always arising which created new problems and for which there must be training. Every man was required to attend school once a week the year around, except for a month's vacation in the summer. The functions of the school were twofold: to give initial training to recruits, and to continue training of the older men.

The new men were put through an intensive 90-day training program with half of each day spent in school studying the performance of police duties and the other half spent on the beat under the supervision of an older officer and a patrol sergeant. After this training was completed the recruits attended school regularly once a week. Teaching was principally by the lecture method and by assigning mimeographed material to be learned and recited. While the regular class work was conducted by the captain of the training school, the various district captains supplemented the instruction. Rules were discussed and material which had a current or special interest was brought to the attention of the officers. Thus the training was both centralized in the school and decentralized through the captains.

A psychologist was employed in 1935 to introduce into the teaching program techniques developed in the field of adult education. The job of the psychologist was to determine what new material should be included in the training program, and to work out more effective methods of presentation. The officers had been taught exclusively by the lecture method. These lectures, repeated year after year, had lost their vitality and became an ordeal to the men who often came to school after 8 hours of duty. Improvement in the method of teaching, therefore, was of vital importance.

THE LESSON PLAN

The lecture method was abandoned, except in those instances where experts were called in to lecture to the school as a whole. A lesson plan was devised in which the instructor did a minimum amount of talking and the class participated in discussion. Techniques stimulating learning, developed in psychological and educational laboratories, were used, including presentation of material visually with slides and moving pictures, utilization of problems pertinent to everyday work, conducting experiments before the class—such as perception of occupied versus unoccupied time, and terminating each study unit with an objective examination. Instruction was planned according to the wide differences in age, education, experience, and intelligence of the officers, e.g., mimeographed summaries of the class discussion were written in simple style and technical material was made strictly applicable to police problems.

Often material was included which showed the officers their own limitations and thus kept them on guard in making statements and ensured the use of the memorandum book for recording facts rather than trusting to memory, e.g., a film called "The Hit and Run Driver" was presented to the class and followed by a true-false test on the facts presented. In an attempt to bring home to the traffic officers the need for consistency in their signals of traffic control, moving pictures were taken of the officers at work, shown to the class, and discussed. After the new training methods had been in practice for a year, an attitude survey was conducted with the purpose of promoting interest in the school and making the officers feel that its function was to give them the training they needed to prepare them for advancement. The results showed that the officers appreciated the new method, and they made many valuable suggestions for improvements.

TRAINING SUPERVISORS

Supervisory personnel included the sergeants, who are the first point of contact between the administration and the regular officers, and the lieutenants and captains, who are responsible for making policies. The sergeant interprets the policy of the department to the rank and file officers and is responsible for seeing that this policy is carried out. It was necessary for him to know what was being taught in the school and the school relied upon the sergeant to see to it that the material learned in the school was put into practice. Because of the military discipline of the police department, sergeants were segregated into classes of their own in the school but they were given the same material each week which was discussed in the classes for patrolmen. Their interest and response testified both to their need for this instruction and to the success of the segregation method used.

On the assumption that the best way to interest the captains and lieutenants in the training school program was to ask their assistance in developing training material, a conference program was worked out for their instruction. Typical conference technique was employed (Shellow and Harmon, 1935). These conferences yielded much valuable material. The captains and lieutenants offered concrete examples from their experience illustrating points discussed. They were anxious to give to the training school the wealth of information, experience, and police techniques which they had gathered during their many years of service. The conference plan proved to be the necessary connecting link between the training of the police officer and the personnel administrators of the department.

FOLLOW-UP PROCEDURES

The development of objective criteria to check tests and training procedures is as important as it is difficult. Probst (1931) has developed a rating scale for public service employees which can be adapted for use in any civil service position. It is, of course, necessary to have several supervisors rate the same men, and to have these ratings made periodically. During the patrolman's probationary period in the training program referred to above, monthly follow-up reports were sent out to the sergeant, lieutenant, and captain of the district to which the patrolman reported. These follow-up reports proved to be valuable in helping the new officer through conferences to adjust to his duties and in indicating at an early period in his service behavior which needed to be corrected. Nine categories were rated as follows: aggressiveness and initiative, learning ability, conduct, judgment, temperament, appearance, physical condition, reliability, and activity.

VALUE OF SERVICE

Once the psychologist has become established as a personnel specialist in a police department, and has proved to the men that his interest is not in "getting something on the officer" but in helping each to find his proper niche in the work of the department by determining his fitness for various types of duties, he can make a real contribution to police morale and efficiency. From the point of view of internal efficiency, the psychologist has much to offer—raising the standard for induction into police service, establishing promotional procedure based on merit, transferring accident-prone drivers, and setting up criteria for evaluating services.

The cooperation between the police department and various social agencies is becoming closer each year. Representatives of these agencies, as well as of the schools and clergy, are constantly being asked to assist the Chief in working for state laws and city ordinances which concern public welfare and safety. The psychologist is accustomed to work with social agencies and specialists of various professions including his own. He speaks their language and understands their point of view. He is in an excellent position to interpret the policies of the police department to the community and the needs of the community to the department.

READING REFERENCES

Diehl, H. S., and D. M. Paterson, *A Personal Study of Duluth Policemen*. Minneapolis: Employment Stabilization Research Institute, University of Minnesota, 1933.

Finlinson, J. J., *Police Training Schools*. Chicago: City Manager Yearbook, 1931.

Martin, E. M., "Experiment in New Methods of Selecting Policemen." *National Municipal Review*, 1923, **12**, 671–681.

Office of Education, *Training for Police Service*. Washington, D.C.: The Office, Vocational Division Bulletin No. 197, 1938.

O'Rourke, L. J., "Selection Standards in Training in Crime Detection." *International Association of Chiefs of Police Proceedings*, 1933, **40**, 70–75.

Probst, J. B., *Service Ratings*. Chicago: Bureau of Public Personnel Administration, 1931.

Viteles, M. S., "Psychological Methods in the Selection of Policemen in Europe." *Annals of the American Academy of Political and Social Science*, 1929, **146**, 160–165.

Vollmer, A., *Police and Modern Society*. Berkeley: University of California Press, 1936.

———, *The Selection and Distribution of Police Personnel*. Chicago: City Manager Yearbook, 1931, pp. 140–148.

39. SUBNORMAL WORKERS

THEODORA M. ABEL, PH.D.

Consultant, Institute for Research in Psychotherapy and Columbia Research Project on Contemporary Cultures; Diplomate in Clinical Psychology of the American Board of Examiners in Professional Psychology. Formerly: Director of Research, Manhattan High School Women's Garment Trades; Senior Psychologist, Letchworth Village, New York State Department of Mental Hygiene.

EMILY BURR, PH.D.

Director, Guidance Bureau, Inc., New York; Instructor, Hunter College; Diplomate in Clinical Psychology of the American Board of Examiners in Professional Psychology; President, Consulting Division, American Psychological Association; President, American Rehabilitation Clinic. Formerly: Clinical Psychologist, Clearing House for Mental Defectives, Post-Graduate Hospital and Department of Public Welfare; Director, Vocational Adjustment Bureau, New York; Instructor, Clinical Psychology, Columbia University and Brooklyn College.

Subnormal occupational workers consist chiefly of those whose levels of general intelligence fall roughly between the I.Q. (Intelligence Quotient) scores of 50 and 80 on standardized tests. Individuals with I.Q.'s above 80 compete more easily with normal individuals and do not present the same adjustment problems as do those who fall below 80. With rare exceptions are individuals with I.Q.'s under 50 able to earn a living, and, although they can be selected on the basis of their special abilities as workers in sheltered shops or in industrial schools, they cannot usually effect an occupational adjustment when in competition with normal individuals.

Investigations show that the jobs in which these subnormal individuals are employed are unskilled and semiskilled occupations, such as farmer's helpers, common laborers, and those doing simple household tasks. Among the subnormal only one out of eight males and one out of fourteen females work at jobs above the levels of unskilled labor (Keys and Nathan, 1932).

THE PROBLEM

The type of problems found in vocational guidance of subnormal workers is not unique; it is for the most part similar to that met in guiding others, regardless of their intelligence level, namely, finding out their interests and aspirations, their assets and liabilities in performing various tasks, and their physical fitness for doing work of a particular kind. But workers with subnormal intelligence need special guidance into the vocations, and supervision after they enter them, since as a group, even on unskilled jobs, they are a greater occupational risk. They are slow learners; they find it difficult to secure their first jobs and to make initial adjustments on the jobs; they have trouble in getting along with employers and other employees; and often are forced to change jobs because they are frequently employed in seasonal occupations.

PERSONNEL SELECTION

The first step in the selection of subnormal personnel is the administration of a standardized test of general intelligence. By this means, those with I.Q.'s falling between approximately 50 and 80 can be recognized and placed in training or in an occupation commensurate with their limited intelligence.

INTELLIGENCE AND VOCATIONAL SUCCESS

Because of the many other factors influencing the individual's fitness for a job, there are differences of opinion as to the value of general intelligence ratings in vocational selection. But the best indicator of success in an occupation over and above actual trial on a job is the individual's general intelligence score and the measurement of the efficiency with which he functions. On the whole, workers with an I.Q. above 80 will succeed better on certain jobs than will those of high-grade moron and borderline intelligence (I.Q. 60 to 80), and, in turn, they will be better paid on jobs and make more successful adjustments than will those below this level. On the other hand, in placement work the fact that the less intelligent individual may be better satisfied on a simple, monotonous job, and hence a greater occupational success than a more intelligent worker on the same job, has to be considered (Cowdery, 1932).

OCCUPATIONAL INTELLIGENCE LEVELS

Various attempts have been made to work out minimum or critical mental age levels for success in a given occupation. Beckham (1930) reports that male workers with an M.A. (Mental Age) of eight years can cut hair, shave, and do simple domestic work; those with an M.A. of nine years can be storekeepers under supervision and can do routine work in millinery and dress shops. Minimum M.A. levels for different factory tasks have been determined by Unger and Burr (1931) for clients of the Vocational Adjustment Bureau for Girls in New York City. Girls with an M.A. of six years are able to do packing and simple factory work; those with an M.A. of seven years can assemble parts, act as errand girls, and do examining and pasting; those with an M.A. of eight years can do straight garment machine operating on nondelicate materials, cutting, and folding; and those with an M.A. of nine years can do hand sewing, press operating, and simple numerical or alphabetical filing. Ten years is the minimum age level for simple clerical work and an M.A. of eleven years is the minimum for salesgirls.

JOB SHIFTS

These M.A. levels have to be modified as technological improvements combine jobs. For example, packing often is done by machine and a girl who might have done packing by hand may be expected to wrap products in cellophane. This task requires the use of the small muscles of the fingers and more precision of movement than is usually found among workers of subnormal intelligence. Many employers demand workers who can adjust to two or more tasks, such as packing and assembling. Even though two tasks may be relatively simple, few subnormals find it possible to shift from one type of task to another.

INDIVIDUAL DIFFERENCES

The assigning of a minimum or critical intelligence level for success with one task can be used at best only as a rough estimate of occupational adjustment. It is a guide that holds true on the average, but not necessarily for the individual. Among subnormal as well as normal people, there are many fine differentiations and individual differences that cut across mental levels and help to determine success and failure on the job. Among these differences are physical health, special abilities and disabilities, social competence, efficiency of functioning, and such dynamic aspects of behavior as attitudes, motivation, and emotional stability. An individual of subnormal intelligence who has another handicap, such as an organic heart lesion, deafness, poor eyesight, a spinal curvature, tremor of the fingers, glandular imbalance, or severe malnutrition is an even greater liability than one of normal intelligence with similar handicaps. Consequently, the physical health of each subnormal candidate for a job must always be investigated, and therapy recommended where indicated.

NONVERBAL ABILITIES

Of particular importance in matching the individual subnormal with a job is the determination of the special abilities that can be used to advantage in training and work. Certain subnormals have special capacities far above their mental levels. These are usually nonverbal in character since it is particularly on the verbal-abstract or scholastic level that the subnormal fails. Nonlanguage tests of general intelligence, for example, the Army Performance, Army Beta Test, Bellevue-Wechsler Performance Scale, the Porteus Maze, or the Pintner Non-Language Scale, can aid in selection of special nonverbal abilities. High scores on these tests indicate satisfactory ability to understand and use concrete data. Individuals with low I.Q.'s on a verbal intelligence scale and I.Q.'s of 100 or more on the Pintner Non-Language Scale have been found to do excellent hand-sewing (Abel, 1939).

MOTOR ABILITIES

Subnormals, as do normals, vary considerably in motor speed, dexterity, and coordination. Quite a few tests have been developed for measuring motor skills in specific occupations. For example, the dotting, tapping, and tracing tests of the MacQuarrie Mechanical Ability Scale indicate speed and control of simple movements; the tapping test also gives an estimate of the regularity of rhythmic movements. Because of rhythmic limitations, subnormals are unable to do the teamwork required in the routing of material in a factory.

The Institute of Educational Research Girl's Mechanical Assembly Test devised by Herbert A. Toops, 1923, and revised by Burr and Metcalfe, 1937, consists of seven practical items and reveals ability in simple assembly, sewing, and trimming. Steadiness of control and visual acuity are suggested by successful cutting on the item requiring trimming. A girl can clip threads, cut appliqué, and do similar simple tasks if she can achieve a score of 7 out of 10 on this item.

The O'Connor Finger and Tweezer Dexterity Tests measure ability to make fine muscular movements. A grade of B is the minimum rating for success on fancy icing in the pastry trade and other work in which neatness, precision, and speed are required.

The Treat Paper Folding Test shows facility in two-dimensional work. The speed and steadiness of control with which it can be done under pressure indicate ability to operate an electric power machine, but an even better indication is obtained when its results are combined with the results of the Porteus Maze and the trimming test of the I.E.R. Girls' Mechanical Assembly Scale.

MENTAL EFFICIENCY

In evaluating the mental ability of the subnormal the efficiency variable must be considered, for often an individual of low mental level with good efficiency can adjust satisfactorily, whereas another individual of much higher mental level is unable to adjust because of poor functioning ability (Babcock, 1940). To know this fact is essential if suitable placement is to be made.

Inability to function to capacity may either limit the M.A. score or prevent sustained work at the level of this ability. Consequently, it is important to obtain an efficiency rating, as well as a mental age rating, of subnormal personnel. A low score on the Revised Examination for the Measurement of Efficiency of Mental Functioning (Babcock and Levy, 1940) indicates either too slow or too quick a response, either of which reduces performance level. For example, a girl who is slow in warming up to new tasks or slow in continuous work should not be recommended for such work as hairdressing or making artificial flowers, in which trades speed of fine movements is demanded. Poor motor efficiency usually indicates inability to make an adequate vocational adjustment.

ATTITUDES

Mental stability and emotional control cannot be overlooked in the selection of subnormal personnel. In fact, an individual who is not only of subnormal intelligence but also has an emotional disturbance or a negative attitude toward a work situation is an occupational liability. Certain subnormals are so nervous that they are unable to tolerate the noises from metals in factories. Many are restless and unable to stand or sit still at one task all day. Some subnormals lack alertness and need special prodding to arouse them to the point at which they show an enthusiasm for work; others are overenthusiastic, have a phantasy life, and a drive to carry out work processes far beyond their mental capacity (Bailie, 1931). Some have patience in doing work that they do not like; some never notice or complain about work conditions; others complain freely and indiscriminately, often in order to hide their feelings of insecurity when trying to compete with faster and more intelligent workers. With some, their feelings of insecurity are so strong that they distort simple instructions into censure and reprimand. All of these modes of behavior have to be detected and taken into account when subnormals are placed on jobs.

TECHNIQUES

No one technique can be used for measuring these differences in personality adjustments. Personal data inventories cannot be administered successfully since subnormals are limited in their comprehension of verbal instructions. The Doll Social Maturity Scale (1936) has been found helpful in evaluating the social effectiveness of the subnormal. Individual interviews, both with the candidate for the job and with those who are in a position to know something of his characteristics, supplemented by information from his social history, remain the most useful guides in the vocational counseling of subnormals. The techniques of measurement of the dynamic aspects of behavior, such as the frustration and projection methods, are in the process of standardization. Vocational counseling should always be available to the subnormal whether or not he is established in a job.

ENVIRONMENTAL INFLUENCES

Many limiting factors in the environment prevent or modify one's success in an occupation. Economic conditions influence the subnormal's vocational adjustment and restrict his range of employment. However, during periods of labor shortage, as in the late war, these subnormal individuals are employed more steadily and at a higher wage than they are in normal times.

During a depression, the subnormal is the first to be laid off, and has difficulty in securing and maintaining employment. He must meet greater competition in finding a job than the normal, no matter how simple and monotonous it may be.

The subnormal person is more easily exploited than is the normal, for his judgment and discriminative ability are rarely adequate in a given situation. Seasonal occupations are more likely to affect the subnormal for he is less able to learn a second job as a substitute occupation during the slack season.

Legislation concerning minimum age of employment and hours and wages determine the point at which a worker is considered substandard. States prohibiting adolescents from seeking employment before the age of seventeen give subnormal persons protection over a longer period, allowing him more time for training in a skill and in general competence.

Standards of proficiency and behavior vary from one employer to another and from one job to another. A subnormal may be successful as trimmer of threads in one concern, whereas he may be discharged for incompetence in another. One who is quiet and unobtrusive may be hired by one employer; another employer may turn him down in favor of a more aggressive and lively individual.

National and cultural groups affect particularly the work adjustments of subnormals. In some groups the parents are averse to having their daughters do housework, a job in which the subnormal can be well trained and in which she receives necessary supervision. All parents are pleased when their offspring find commercial rather than industrial jobs, although it is in the latter that subnormals are more likely to achieve a degree of success.

TRAINING

Placement of the subnormal worker in any but the most crude and unskilled jobs is more effective if there is training along the lines suited to the adjustments he must make to be successful in the working world. This is particularly true if he is under the guidance of school counselors who know his needs.

TRAINING PROGRAM

Opinions differ concerning what should comprise a training program for subnormals and how specific or general their vocational and prevocational education should be. The consensus is that subnormal pupils need a specific vocational program, for which they show both interest and ability and in which employment opportunities exist.

If the subnormal pupil is emotionally stable, he usually enjoys learning thoroughly one task. This characteristic should be capitalized upon during a period of vocational training preceding his withdrawal from school. He should be aware that the task he is learning is in preparation for a job. He thus sees purpose in the daily school procedure and has hope of assuming a responsible role in the community when he enters the working world. His vocational instruction should be practical. It should be continued until he attains an acceptable degree of perfection in which rate of production is equal to that of the average worker for the occupation. But his specific training should be so gauged that he reaches maximum speed and quality of production not too long before he is allowed by law to go to work. If the subnormal is ready at fifteen to do a particular job, but is not allowed to be employed before he is sixteen or seventeen, he becomes discontented, forgets the process he has learned, and sees no reason for staying in school another year.

LEARNING PROBLEMS

Since subnormals are slow learners, they usually need considerable time to learn tasks which the normal learn quickly. The slow learner should not be hurried in the initial stages of his training. Only gradually should he be encouraged to speed up. A system of small rewards often is effective in encouraging him to work faster.

In training, the individual differences and idiosyncrasies of subnormals have to be given more consideration than do those of normals. Subnormals sometimes have trouble starting new tasks; some of them progress just so far and then stand still, or even regress in the quality of their work; others have to be shown step by step how to do very simple tasks. Only trained and experienced teachers can effectively help the subnormal to overcome the difficulties he encounters while learning a skill.

SOCIAL EFFECTIVENESS

Subnormals require special training in matters pertaining to social competence—personal hygiene, handling money, courtesy, punctuality, and the like. They require special drill in reading, so that they can achieve at least a third- or perhaps a fourth-grade level of proficiency. Training in creative effort—embroidery, crocheting, music, dramatics, or woodcarving—is desirable to help them find an activity suitable for their leisure time.

WORK ADJUSTMENT

In spite of careful vocational guidance and specialized training, there can be no complete certainty of the degree of adjustment that will be attained by the subnormal in his actual work situation. Working conditions for him are quite different from the school or test situation. Various methods have been suggested or tried out more or less successfully for making initial adjustments on jobs by subnormals which includes apprenticeships, learner's permits giving temporary exemption from minimum wage requirements, exemptions from minimum wage standards, sheltered shops, cooperation of employers or trade unions with schools in teaching specific work processes, and training courses under special guidance in a vocational school or an actual factory setting. Unfortunately, none of these has been employed consistently over a period of years or extensively enough to handle the number of subnormal individuals ready and willing to enter the working world.

IDLENESS

Perhaps the most difficult obstacle for the subnormal to successful work adjustment is a long period of idleness while waiting for the initial job, during a slack season, or after he has failed on a job. Waiting is particularly detrimental when the subnormal has learned a skill requiring continual practice in order to maintain speed and efficiency. Except in rare instances, the subnormal who has discontinued his schooling usually has no place to go to keep up necessary practice.

OUTLOOK

Only large-scale cooperative effort on the part of education, social welfare agencies, government, and organized labor can work out an effective program for providing subnormals with the training they need in school prior to going to work, preceding their initial and subsequent adjustments in the working world, and to tide them over periods of idleness. If this bridge could be crossed, the subnormal would have a good chance to become an efficient and steady worker. He will stick to his work because he is willing to accept a routine, monotonous job and has no aspiration to advance to superior positions.

READING REFERENCES

Abel, T. M., "A Study of a Group of Subnormal Girls Successfully Adjusted in Industry and the Community." *American Journal of Mental Deficiency,* 1940, **45,** 66–72.

———, and E. F. Kinder, *The Subnormal Adolescent Girl.* New York: Columbia University Press, 1942, Chap. 3; Chap. 4.

Anderson, M. L., "The Education and Training of Older Dull Pupils." *Proceedings and Addresses of the American Association for the Study of the Feeble-minded,* 1931, **36,** 100–109.

Babcock, H., "Personality and Efficiency of Mental Functioning." *American Journal of Orthopsychiatry,* 1940, **10,** 527–531.

———, *Time and the Mind.* Cambridge, Mass.: Sci-Art Publishers, 1941, pp. 145–149.

Beckham, A. S., "Minimum Intelligence Levels for Several Occupations." *Personnel Journal,* 1930, **9,** 309–313.

Burr, E. T., "The Vocational Adjustment of Mental Defectives." *Psychological Clinic,* 1931, **20,** 55–64.

Channing, A., *Employment of Mentally Deficient Boys and Girls.* Washington, D.C.: US Department of Labor, Children's Bureau, Publication No. 210, 1932.

Doll, E. A., "Education and Training of the Feebleminded." *Proceedings and Addresses of the American Association for the Study of the Feeble-minded,* 1932, **37,** 99–116.

Keys, N., and J. M. Nathan, "Occupations for the Mentally Handicapped." *Journal of Applied Psychology,* 1932, **16,** 497–511.

Metcalfe, Z., and E. T. Burr, "A Practical Form of the Girls' Mechanical Assembly Test." *Journal of Applied Psychology,* 1936, **20,** 672–679.

Unger, E. W., and E. T. Burr, *Minimum Mental Age Levels of Accomplishment.* Albany: University of the State of New York, 1931.

40. EMPLOYMENT AGENCIES

LORIN A. THOMPSON, PH.D.

Director, Bureau of Population and Economic Research, University of Virginia. Formerly: Director of Population Study, Virginia State Planning Board, and of Regional Department of Economic Security, Cincinnati, Ohio; Head, Department of Psychology, Ohio Wesleyan University.

The personnel department of a business or industrial firm is concerned chiefly with the selection of applicants for specific jobs in which the specifications are reasonably well known. The task of the employment agency is more varied and complex since it must measure and appraise the ability, training, and experience of applicants in such a way that the classifications made will meet the specific requirements of different employers. In this respect the work of an employment agency is similar to a vocational guidance clinic. The methods used by persons in finding jobs are many and varied. Firms without personnel departments usually delegate the responsibility of hiring new employees to foremen, department heads, or managers. In such instances the ability of the applicant to sell himself is one of the chief aspects of the employment process. The advantages of a personnel department or an employment agency in selection and placement are that the employee has a wider choice of jobs and the employer a larger choice of applicants.

Employment agencies have changed greatly during the past forty years. Early in this century laborers were often recruited in larger cities by labor brokers, given a one-way railroad ticket to the job. Skilled workers and employers often used the labor unions as the job mart. During the early part of the century private employment agencies, many of which specialized in teacher placement and professional personnel, emerged. As industrial concerns increased their personnel, some installed employment offices. These offices later became personnel departments and in the more modern type of organization have been enlarged into industrial relations departments dealing not only with selection and placement, but with promotion and advancement of workers and the negotiation of employee agreements. Public employment agencies were established in a few cities about the turn of the century. The United States Employment Service was established by an Act of Congress in 1917. It functioned during World War I effectively. After the war its activities were restricted and it did not come into prominence again until after 1930. The passage in 1934 of the Wagner-Peyser Act broadened its activities and stimulated the establishment of state employment services which since have been affiliated with the United States Employment Service. During World War II the State Employment Services were federalized and administered by the United States Employment Service. As a part of the War Manpower Commission the public employment service played an important role in recruiting workers of all kinds for war industries.

JOB REQUIREMENTS

The specifications of similar jobs in different firms have many common characteristics. They are most apt to differ as the result of personal preferences of individual employers. If men and women are to be successfully matched with job requirements the skills and personal traits required by each individual firm should be known. This calls for a reasonably accurate job analysis in which the skills, abilities, and traits are stated in quantitative terms. If this is done by each firm, the task of the employment agency is defined more sharply. One of the ever-present difficulties, however, in matching people with jobs is the fact that the characteristics of the job are seldom defined in the same terms as employment offices use to measure the capacity, experience, skill, and promise of the applicant. For example, intelligence tests have been effective in classifying clerical workers by the employment office. The employer, however, may specify a clerk's requirements in such descriptive terms as accuracy, speed, steadiness, and dependability. There is a real need for a common language in which employer and employment office mutually agree on the meaning of their respective terminologies.

TESTED REQUIREMENTS

A machinist has at least four essential qualifications: (1) the ability to read and understand directions and instructions, (2) the ability to interpret and use drawings and patterns, (3) manual dexterity, and (4) manipulative skill. The employment office can classify "would-be" machinists from "real" machinists fairly successfully by using carefully selected tests. If the employer states his needs in terms similar to those used by the employment agency, the task of referral and placement is greatly facilitated.

To classify machinists on the four characteristics described above the following program may be used. The ability to understand directions can be measured by a number of the more common intelligence tests, such as the Otis Self Administering Higher or Intermediate Examination. Ability to read blueprints can be measured by a blueprint reading test. Many such tests have been devised by business firms to meet their own specific needs. If this procedure is followed it is important to standardize the problems and the scoring. The third and fourth requirements can be met by various manual dexterity and manipulative tests. The Minnesota Rate of Manipulation Test can be used to measure skill and movement and the Minnesota Spatial Relations Test to measure comprehension of patterns and spatial discrimination. The Kent-Shadow Industrial Form Board can also be used to measure pattern and form relations and the O'Connor Finger and Tweezer Dexterity Tests to measure manipulative skill which involves finesse of movement. This latter test was restandardized and new norms developed at the Cincinnati Employment Center, pliers being used instead of tweezers since a Crescent No. 1033 plier is used in many of the manufacturing assembly operations.

The pattern of the applicant's performance on this battery of tests indicates how well he fits the job requirements of the individual firm. Most jobs require a minimum performance level with respect to essential functions. Selection procedure at the employment agency must therefore be designed to meet this requirement. A candidate who falls below the minimum level of performance on any essential characteristics of the job is disqualified.

NONTESTED REQUIREMENTS

Such characteristics as dependability, cooperation, and integrity are important requisites of workers. Rating scales, methods of evaluating previous experience, and personal inventories have proved useful in indicating the value of these qualities in relation to the job requirements. There seems to be no satisfactory method of quantifying the results of letters of reference. Such judgment is difficult to quantify, but despite this limitation letters of recommendation should not be ignored in the general appraisal of the client.

CLASSIFICATION

The task of classifying applicants in an employment office consists of the determination of their qualifications (1) for their primary occupation and (2) for alternate occupations. Such a procedure is equally important in times of prosperity and depression. Many workers find that where there is relatively little demand for them in their primary occupation in times of depression there may be a demand in other occupations in which they can successfully engage. During the period from 1935 to 1938 the Cincinnati Employment Center used a Multiple Occupational Classification plan. Stated simply, Multiple Occupational Classification lists two or more occupations for a person if his ability, experience, and skill justify it. The plan used in Cincinnati was to use a separate punch card for each occupation. The first part of each card contained the same personal data. The second part contained the appropriate occupation and industry codes and the interviewer's rating of the applicant's fitness for such occupation. This procedure enables the employment office to build up a comprehensive and complete inventory of skills. It also increases a person's chances of securing a job, since consideration is not restricted to a single occupation. In Cincinnati more than 50 per cent of the placements made by the Employment Center were from alternate occupation cards. In 1940–41 the United States Employment Service found it desirable to broaden its classification to two or more occupations as a means of discovering workers for essential war work. Most people can work successfully at more than one occupation if need be. A multiple occupational classification plan is most essential from the point of view of helping people find jobs when there are more people available for work than there are jobs to be filled and in utilizing manpower to the best advantage when there is a labor shortage.

Clients or registrants are classified by the employment agency according to occupation. The United States Employment Service Dictionary of Occupations describes briefly the essential activities or skills of approximately 35,000 different listed occupations. There are in addition codes for the classification of jobs by industry and occupation which are used by most federal agencies, such as the Bureau of the Census, the Bureau of Agricultural Economics, the Bureau of Labor Statistics, and the Social Security Board. These classification systems are helpful in standardizing terminology, but are much too elaborate for an employment office.

In order to function efficiently the employment agency needs to study and analyze the specifications on its job requests and translate the requirements for various jobs into terms similar to those used in the classification of the applicant. Such personal characteristics as poise, pleasantness, and initiative are not always indispensable requirements for successfully performing the technical aspects of the job. Moreover the pattern of personal traits desired by different employers varies. Women working as models in fashion establishments are chosen on the basis of height, weight, build, carriage, manner, and attitude, in addition to at least average mental ability. Important special requirements of telephone operators, in addition to intelligence, alertness, and speed are courtesy and a pleasant voice. A supervisor or foreman, in addition to having above average intelligence, must know the jobs which he is to supervise, and have such personal qualities as understanding, patience, and fairness in order to deal successfully with those who work with him. The feeling on the part of the worker as to whether he works *for* or *with* the boss is of major significance in encouraging an atmosphere of loyalty, good will, and efficient work. Thus, job classification in an employment agency involves a careful classification of clients with respect to all common characteristics required by employers and to certain essential qualifications of individual employers.

WORKER ANALYSIS

Testing in an employment office is most effective if the tests are designed to measure job characteristics. But often the job descriptions of business firms do not afford a satisfactory basis for determining the essential characteristics and traits required for satisfactory performance on the job. In this case the psychologist in the employment office translates the loose general descriptions into characteristics susceptible of measurement and chooses such tests as will best meet these requirements. In doing this it is essential to separate the indispensable characteristics of a job from the unusual or specialized specifications of the employer. Such unusual requirements are usually a matter of personal preference. Insofar as job requirements have common characteristics, the norms for the tests used should be standardized on many registrants in the employment agency. Additional special tests are administered to clients as the special requirements vary between the job requirements of different employers.

Job order forms provide a basis for establishing and revising the patterns of traits and abilities for various kinds of jobs. The job order form is frequently a part of the employer's file. When properly kept it serves an important research purpose in addition to its indispensable use in the office.

TRADE TESTS

The classification of domestics and skilled, semiskilled, and unskilled workers, according to tested achievement is necessary. Workers in the skilled occupations often object to written tests but they usually are quite willing to take oral and performance trade tests. Thompson and others (1936) of the Cincinnati Employment Center prepared a manual of oral trade questions to be used by interviewers in appraising applicants who wished to register as skilled workmen. These questions of trade knowledge are a revision and extension of the Army Trade Tests (Chapman, 1921). A facile talker with a general fund of information does not often deceive the interviewer administering a trade test because the questions covering the more exact and technical aspects of the job serve as important clues in detecting a bluffer. Trade questions are effective aids in classifying skilled workers. The Division of Standards and Research in the United States Employment Service has prepared more comprehensive sets of trade questions for use in the public employment offices.

APTITUDE TESTING

The employment interviewer often must determine the fitness of applicants for jobs in which they have had little or no experience, e.g., in selecting workers for aircraft factories, shell-loading plants, and munitions plants. To meet the exigencies of this situation, aptitude tests may be employed, such as those mentioned in measurement of manual dexterity and manipulative skill of machinists. Public employment offices have found batteries of those tests of great value in selecting for apprentice training programs. Such tests eliminate applicants with such characteristics as poor space perception, slow eye-hand coordination, clumsy manipulation, and awkwardness. Wood and Long (1940) devised a number of simple construction tests to indicate aptitude for machinists, carpenters, bricklayers, and machine operators, which were used in the Employment Consultation Service in Cincinnati. This office has shown how talent can be discovered among older workers as well as younger ones through aptitude measurement. Many people have undiscovered aptitudes which if trained can be put to good use both for themselves and for society. Similar procedures were used by the United States Employment Service during World War II in recruiting trainees for the many specialized war jobs.

INTERVIEWING

There are workers who meet the educational and technical requirements of jobs but lack the necessary combination of personal qualities. For this reason the selection of executives usually involves different methods from those for skilled workers. It is important to recognize in the employment agency that personal requirements as well as technical and educational requirements may vary considerably for the same job from one organization to another. Personal adaptability often makes the difference between success and failure on a job.

The personal requirements of applicants are estimated in the interview. It is found that if the interviewer will attempt to determine what the individual might do under conditions of embarrassment, strain, or excessive responsibility, he will secure a more satisfactory appraisal of the client's personal qualities. Such characteristics as talking too much, being impatient, and lacking discretion in difficult situations may completely disqualify an otherwise high grade technical person.

An estimation of personal qualities can be gained from an analysis of the experience record of the applicant for indications of the way he has dealt with difficult and complex situations in the past. The selection of men who are to be entrusted with responsible positions of supervision requires the appraisal of the more elusive psychological factors which make for emotional balance, stability, and social acceptability.

USE OF MEASURES

An applicant considered for placement must meet successfully each of the minimum qualifications of the employer's job requirements. His standing in relation to specific job requirements is shown by a profile drawn up for him from assembled test scores and interview estimates. This profile indicates his standing in each trait in relation to norms established on successful groups of workers. The common practice of averaging test scores violates essential requirements of test interpretations.

Low standing in a single qualification may disqualify a person for consideration in a particular job. Persons working for the federal government in civil service jobs must be American citizens and if this condition is not met the applicant cannot be considered. Other individual qualifications are equally important in various jobs. Keen color discrimination is essential in certain manufacturing operations. If a person happens to be color blind, he is completely debarred because of color deficiency, his other qualifications being unusable for this reason.

SOCIAL VALUE

Psychological methods of interviewing and testing by employment agencies, either public or private, not only facilitates the adjustment of individuals to firms for whom they work, but can contribute to their broader personal adjustment. Tests help select persons for jobs and for transfer from one type of work to another where more accurate means than length of experience are needed for selection. Measures of aptitude are used where there is a scarcity of experienced workers. Testing provides a means of discovering the hidden reservoirs of talent which are available in the community and in the nation. The refinement of testing methods now used to meet the special conditions of transferral of workers from one job to another where related experience contributes essential training has proved of great social value in a national emergency.

Psychological personnel methods have done much to enlighten managers and people generally regarding the human contribution to industrial production. They have emphasized the importance of intangible psychological qualities contributing to industrial cooperation and adaptability of individuals in the occupations. The social responsibility of the personnel officer has been clearly defined by these psychological methods. The employment agency assumes equal responsibility with him in the scientific placement of workers. It must use the best methods to reduce social wastage; in fact it can accept leadership in this as a source of information for personnel officers.

READING REFERENCES

Burtt, H. E., *Principles of Employment Psychology* (rev. ed.). New York: Harper & Brothers, 1942.

Chapman, J. Crosby, and D. R. Chapman, *Trade Tests.* New York: Henry Holt and Company, Inc., 1921.

Münsterberg, Hugo, *Psychology and Industrial Efficiency.* Boston: Houghton Mifflin Company, 1913.

Stead, W. H., et al., *Occupational Counseling Techniques.* New York: American Book Company, 1940.

Thompson, L. A., et al., *Interview Aids and Trade Questions for Employment Offices.* New York: Harper & Brothers, 1936.

Viteles, M. S., *Industrial Psychology.* New York: W. W. Norton & Company, Inc., 1932.

CHAPTER VI

INDUSTRIAL PSYCHOLOGY

41. Worker and Manager	Daniel Starch
42. Efficiency in Work	Rexford Hersey
43. Occupational Safety	Gilbert E. Teal
44. Building Illumination	Miles A. Tinker
45. Industrial Ventilation	L. P. Herrington
46. Industrial Morale	Francis F. Bradshaw and Herbert E. Krugman
47. Military Morale Procedures	Francis Keppel
48. Color Vision	Dean Farnsworth
49. Bio-Mechanics	Jack W. Dunlap

41. WORKER AND MANAGER

DANIEL STARCH, PH.D.

President of Daniel Starch and Staff; Consultant in Applied Psychology and Business Research. Formerly: Professor of Psychology, University of Wisconsin and Harvard Graduate School of Business Administration.

Viewed comprehensively, industrial and vocational psychology include: (1) selection of personnel, (2) training of personnel, (3) periodical evaluation of personnel, (4) job specifications, (5) job evaluation, (6) direction and supervision of personnel, (7) morale and welfare, (8) labor relations and collective bargaining, (9) relations with colleagues and superiors, (10) dealings with customers and the public, and (11) dealings with suppliers, financial houses, and security holders. This chapter will be limited primarily to items 6, 7, and 8 above, that is, the personnel and managerial relations in industry.

THE PRACTICAL PROBLEM

Business is about 85 per cent men, 10 per cent materials, and 5 per cent money. The price paid for products is about 85 per cent the result of the work of men. Raw materials untouched by man exist only in the ground, in the sea, in the forest, and in the air. An apple in the market may appear to be raw material to cooks, canners, and bakers, but it is a finished product in which the brain and hand of man have played a large part. To be considered raw material the virgin, untilled soil and seed from a wild apple tree would have to be taken into consideration. The rest of the value of a finished product is the result of man's labor and the management that directed that labor. The value added by work to the soil and the seed to produce the apple pie on your table is probably at least 85 per cent the result of mental and physical work of people.

Eighty-five per cent or more of the producing assets of a business are human assets. The managers and employees of a business are worth approximately fifteen times as much as its physical assets. A business capitalized at one million dollars usually employs 250 persons since it takes an investment of $4,000 on the average to employ one person. The average employee is paid $1,800 a year. Capitalized at 3 per cent that pay of $1,800 represents an investment of $60,000 per employee. Hence, the total working capacity of 250 employees is an asset of 15 million dollars as compared with the one million dollars of physical assets carried on the books of the company.

The causes of failures in business are 80 per cent within the control of management and worker. Analysis of numerous failures shows that the factors associated with them, such as competition, insufficient capital, adverse domestic and personal factors, dishonesty and fraud, bad debt losses, inadequate records, are due to poor judgment and incompetency among managers and workers.

Nearly all business problems reduce to human problems. Nothing happens in business except through people. The average executive of necessity devotes almost all of his time to dealing with men—directing, supervising, persuading, and planning. Production problems, distribution problems, financial problems, and administrative problems nearly all end up as human psychological problems.

A Management Illustration. These facts are clearly illustrated by the tale of two men who came to a banker for a loan. They needed $150,000. It was purely a financial problem they said. The banker knew the firm as an oldtimer in its industry. The two men told how the business had been founded by their two grandfathers, carried on by their fathers, and had now been in their hands for approximately 20 years. In grandfathers' time it had led its industry. In fact, it would lead its industry today again if only they had the capital to operate on, they said. That is why they needed $150,000. Further inquiry brought out the fact that 25 years ago the chief competitor who now leads the industry was a mechanic who started up on his own with little money. His business came up while theirs went down. The second generation then in charge of their business, the fathers of the present managers, took long vacations, withdrew money freely and spent it liberally. Their sons came into a business that had lost its momentum. They lacked basic experience; they had not even their fathers' advantage of being trained by the two hardfisted old founders. They had expensive tastes. They paid themselves excessive salaries, did not attend to business, and dissipated the company's money by paying dividends even when there were no earnings. With this information in hand the banker refused to make the loan and then said: "You haven't a financial problem; you have a human problem. Maybe you can pull yourselves together and rebuild your business; you have a lot better start than your competitor in competition against your company in its heyday. If you can't do it for yourselves, probably you can find someone competent to run the business as it must be run to make money."

Management is largely management of men and therefore falls within the domain of psychology. Hence, psychology can contribute an important part to the training of executives. To this end psychologists and managers should recognize that they must study, analyze, and investigate management problems by scientific methods.

Attitude of workers is tremendously important. Both manager and worker have the responsibility of cultivating work attitude. It may easily make a difference of 25 per cent in production.

A Worker Illustration. Western Electric Company began in 1924 a most significant experiment in human relations. It was undertaken to determine the effect of various factors on performance and output of workers. Two groups of employees were selected, a test group whose working conditions were changed, and a control group whose conditions remained the same.

First, the test group was given increased light. Its output went up. So did that of the control group with no change in lighting. Then the light of the test group was decreased below that of the control group. Its output went up again, and so did that of the control group. A dozen variations were made with the test group, such as rest pauses of varying lengths during each half-day, hot snacks, lengthening or shortening the working day, Saturdays off, and last, removal of all changes and return to the same physical conditions as at the beginning. Output held even or went up irrespective of changes, whether adverse or favorable.

The results were mystifying. Evidently an unknown factor had thrust itself into the experiment. It was the changed human attitude, the way the workers felt about their work and their group. They had been asked to help solve a problem. This changed them from insignificant cogs in a machine to a congenial, important group. The workers in the experiment had been placed in a special room. An observer made note of everything significant that happened. He told them about the experiment and was their counselor and friend, inviting comments and listening to complaints. Their output increased from 2,400 units a week before the experiment to 3,000 units, a gain of 25 per cent, in the last test when they were working under the same physical conditions as at the beginning.

THE PSYCHOLOGICAL PROBLEM

What should managers know about human nature in order to deal intelligently and effectively with people? What should workers know? The fault of executives most frequently mentioned by employees centers around one concept, namely, failure to understand and consider the other fellow. In other words, the commonest fault of executives is unfairness. This statement is based on confidential, personal interviews with one hundred and fifteen employees and junior executives in business organizations. Sixty-eight per cent of these persons expressed this fact in one way or another, such as, "he treats employees as working for instead of with executives," "lacks consideration of associates and subordinates," "not willing to listen to subordinate's point of view," "orders employees instead of requesting them," "makes decisions without regard to how they affect the persons concerned," "grievances are not listened to," "lines of authority are not followed," "men doing less important work receive as much as those doing more important work."

By the same token the chief fault of workers is failure to recognize the problems and point of view of managers, and to realize that ultimately the social and economic objective of both is the same. Workers are woefully uninformed about the elementary facts regarding the very business they are employed by. For example, according to a recent poll railroad employees thought that railroads earned 27 per cent on their investment. Actually they earned only 3 per cent in 1945. According to another poll workers believed that they receive 25 cents in wages for every 75 cents that went to management and stockholders, that companies were making profits of 20 to 25 per cent and that 10 per cent would be a fair maximum. Actually business earned less than 6 per cent on net worth during the period of 1927 to 1943.

MOTIVATION

Essential for the effective handling of the problems suggested by these comments is a thorough understanding by managers and workers of motivation in human nature, that is, what makes us think and feel and, particularly, act as we do. Sixty-seven noted psychologists, including almost all the living past presidents of the American Psychological Association answered this question: "Considering our general environment, that is, living and working with people, and considering one's own happiness, what are the most essential things to know about human nature or people?" Forty-six of the sixty-seven psychologists, or two thirds, gave first place to motivation. The relative importance attached to this aspect of human nature is indicated by the number of times other factors were mentioned by these psychologists: motivation, 46 or 68 per cent; emotions, 19 or 28 per cent; habit formation, 18 or 27 per cent; personality, 10 or 15 per cent; thinking and imagination, 10 or 15 per cent; social adjustment, 10 or 15 per cent; effect of environment on behavior, 10 or 15 per cent; intelligence, 9 or 13 per cent; and how to evaluate individuals, 8 or 12 per cent.

The manager or worker must understand human motivation if he would have at his command the means to influence and direct others. He can get others to do what he wants by compulsion and suasion. Compulsion is the method of slavery; suasion is the method of free men, and the method by which motivation in workers and managers must be gained to achieve results in business and industry. Two laws of motivation in which suasion is grounded can be designated as the Law of Egocentrism and the Law of Induction.

LAW OF EGOCENTRISM

Each person is the center of his world of experience and action. You are the center of your world and I of mine. All experiences are seen from that center and all actions are designed to serve that center. Each person is to himself the most important object in the world. This is egocentrism; it is a basic law of motivation.

Scholars, thinkers, and men of affairs, since time immemorial, have recognized that man was egocentric. This *is human nature*. "Self-preservation is the first law of nature," said Samuel Butler and Diogenes put it, "An animal's first impulse is self-preservation." But the law of egocentrism is so obvious that employer and employee in dealing with each other do not see it, like the glasses on the forehead of which we are not aware. Not being conscious of it, we fail to observe it in directing others.

Our thinking and our language are completely infused with egocentric expression. Every fifth word we speak or write is *I, me, my, we,* or *ours,* and every tenth word is *you* or *yours*. *You* and *I* with their variations are by far the two most frequently used words. In a count of 5,136,816 running words in personal and business letters and in miscellaneous writing by various investigators, Horn (1927) found that *I,* with its variations, occurred 1,239,338 times and *you,* with its variations, occurred 520,652 times. Your photograph is to you the most interesting picture. The camera industry would collapse if "you" and "I" were removed from all pictures. A man buys a flashlight photograph of a banquet for the satisfaction of locating himself in it.

COROLLARY 1. The law of egocentrism has two corollaries of application. Since each person is the center of his world of experience, cause him to feel that he *is* the center. This is corollary one. Man's deepest wish and inherent right are to be regarded as a person among persons and to be so treated by his fellowmen. Whiting Williams (1925), who studied, worked and lived with thousands of laboring men to find what was on their minds, concluded that the strongest desire is the wish for worth. In a letter to his students in Radcliffe College who had presented him with a beautiful azalea plant, William James, America's greatest psychologist, wrote: "I am deeply touched by your remembrance . . . I now perceive one immense omission in my Psychology, the deepest principle of Human Nature is the craving to be appreciated, and I left it out altogether from my book, because I had never had it gratified till now . . . ."

COROLLARY 2. Since each person's behavior is designed to serve himself, cause him to feel that what you wish him to do will serve him. This is the essence of the psychology of selling as manager, advertiser, or labor leader. People will not do what they are told merely because they are told (except under conditions of compulsion). What they are told must serve their self-interest to bring action. An illustration that will make this clear follows:

> Banks failed everywhere during the depression in 1932. Depositors became frightened and withdrew their money by the millions. President Hoover appointed a committee headed by Colonel Knox to quiet fears and stop runs on banks. Depositors were told not to fear; the banks were sound. There was a slight halting for a few weeks. Then withdrawals increased and shot up to unprecedented heights. Despite all they were told, depositors became increasingly fearful. Money, they felt, was safer in a mattress than in a bank. Only the proc-

lamation of a bank holiday by President Roosevelt immediately after his accession to office finally stopped the vicious circle.

LAW OF INDUCTION

Behavior of one person induces corresponding behavior in others. Look intently at the sky for one minute and almost everyone who sees you will do likewise. Three short paragraphs were given to each of one hundred and six adults in an experiment by Starch (1910). One paragraph was in typewritten form, the second in vertical script, and the third in slanting script. The instructions were to write in their usual manner the three paragraphs before them. The degree of inclination of all tall letters, such as *H, F, G* and *L,* was measured. The handwriting of these one hundred and six persons was on the average 7.3 degrees nearer to vertical when they had the vertical sample before them than when they had the slanting sample before them. The degree of slant of writing produced with the typed paragraph before them was between the slant of the other two. Handwriting in adults is a firmly fixed habit. Yet samples on paper before them had a material effect on this established habit. The active behavior of a dynamic person will produce a far more powerful effect on the actions of others.

FADS AND FASHIONS

Thousands of persons bought land in Florida at from 10 to 20 times normal prices because thousands of others had done so. Millions of persons paid from 10 to 20 times fair prices for stocks in 1929 because millions of others had done so. Many times normal prices were paid for tulips in 1634 during the tulip craze in Holland because everyone was doing so. One bulb was sold for $5,200. Fashions go in waves lasting from a year or two to a generation or two. Miniature hats were worn by women during the Civil War; they were again worn in 1940. It was the fashion for several decades about 1800 for men to shave. Then for several decades about 1860, it was the fashion to wear beards. In about 1870 someone invented a safety razor, but it was a failure because men then wore beards. About 1900 when the safety razor was reinvented and men no longer wore beards, safety razors were instantly successful. The living behavior of men and women is a subtle, potent influence on the acts of others. The law of induction must be understood and observed by him who wishes to lead and direct rather than drive and command.

APPLICATION OF LAW

The law of induction may be paraphrased in the office and the shop in the phrase: as the boss is, so is the gang. For the executive and supervisor it says: behave and act as you would have others behave and act. For the worker it says likewise: behave and act as you would have your fellow workers behave and act. The successful leader or manager or worker must exemplify in his behavior the qualities of poise, fairness, truthfulness, sincerity, cheerfulness, perseverance, and thoroughness, to mention a few of the most essential ones.

The only way by which one human being can properly influence another except by force or example is by encouraging him to think for himself. This statement applies to both worker and manager. The rules to guide thinking, which guides action, are to state your objective; suggest ideas, solutions, and lines of thought. Skillfully guided in thinking, the person will persuade himself. Don't assert; instead, ask questions. A man will do his own thinking. Step by step he will arrive at your answer. But he will consider it his answer. He will then hold it to be sound and act on it.

TECHNIQUE

How can the manager, supervisor, and worker use these laws of motivation? There are numberless ways. As typical of those of wide application, the following may be mentioned:

a. Show a person the respect due him as a person.
b. Show confidence in him. Give him tasks that are at the top of his reach. Give him responsibilities.
c. Ask him to help you solve his problems and yours. "If you would make a person a friend, ask a favor of him," said Benjamin Franklin.
d. Show him appreciation.

A prominent executive makes a point of showing appreciation at periodic intervals to his immediate associates and subordinates because of the lift it gives them. Experiments show that simple, direct, approving, unhurried, and positive directions to children are decidedly more effective than cajoling, commanding, and threatening directions. Pleasant and positive requests, such as, "Please do," are more effective than the scolding type, "I don't see why you refuse," or the like.

An Illustration. A gruff executive sent a young man to the personnel officer with the complaint that he was intelligent and a good worker but so conceited nobody could stand him. The executive had given him a "dressing down" every day, but the more he took him down the worse the employee got. The story of the young man's life was a clue to his braggadocio. Behind him lay a broken home with a sense of deprivation and resentment. His resulting feeling of inferiority led to boasting to build up his shaken self-confidence. The personnel director changed the boy's job slightly so that he was given more definite responsibility. He reported to the executive and not to a woman subordinate, as had formerly been the case. He responded instantly with approved behavior as there was no longer any need for him to boast in order to sustain his confidence. He became a useful, efficient, and agreeable employee. But a more striking transformation was the changed attitude of the executive. He was so impressed by this apparently deep insight into human behavior that he was more improved than the young man.

e. Ask a person how he would discipline himself when he has failed to do or perform as he should. If you do so calmly and sincerely, he will be likely to prescribe a more severe corrective than you would yourself.

f. Help him save his self-respect. If he has made a mistake, make it easy for him to admit it and save his face.

An Illustration. To conserve wheat during World War I, all merchants handling flour were required to sell an equal quantity of substitutes, rye, buckwheat, etc., with each flour sale. But merchants generally were uncooperative. Trouble-shooters by the dozen were sent out from Washington to keep merchants in line. But there were not enough of them. Finally, Walter Innes, Food Administrator for Kansas, sent to a merchant, who was reported to be breaking the flour regulations, the following wire: "You are hereby appointed representative of the United States Food Administration and directed to call a meeting of all merchants in your town advising at once whether they are all on the 50-50 basis selling an equal quantity of wheat substitutes with each sale of flour. Walter Innes, United States Food Administrator for Kansas." Proud of the tribute to his leadership, the offending merchant promptly revised his own practices and saw that his fellow merchants did likewise. This method worked so well that Herbert Hoover referred the telegram to other state administrators. Often an "All O.K." reply came back the same day.

g. Listen to a person if he has troubles or complaints. Let him talk himself out. If you stop him in the middle of his criticisms and try to answer him, you merely stir his fire of emotions and make him more severe and unfair. When he is talked out, his fire will be burned out. He will then be calm and receptive to what you say.

h. Admit your own mistakes if they affect him. Whether you are an executive over a thousand persons, a supervisor of one helper, or an individual worker dealing with your superiors, admitting your own mistakes will restore the other fellow's respect for you.

i. Do him kindnesses; do so with sincerity and appropriateness.

It is not sufficient to recognize the laws of motivation in the abstract. It is absolutely essential to hold them before one's eyes consciously and avowedly and study their continual application in all that one does. They apply in all human relations, whether manager, worker, salesman, or consumer. The physicist may understand in all detail the law of gravitation but unless he is also the practical engineer he will not be able to build a bridge that will hold. Psychologists, and both managers and workers may be good psychologists, may thoroughly understand the laws of motivation, but to profit from this knowledge there must be planned, intentional application.

THE OUTLOOK

Manager and worker psychology is an untilled but fertile field. The possibilities of service are unlimited. Its effects on business operations have wide reaches. The approach is twofold: through specific research on management and labor problems by cooperation between psychologists and managers or psychologists and labor groups; and through advisory consultation by psychologists with managers and worker groups on the practical handling of specific problems and situations.

READING REFERENCES

Magoun, F. A., *Problems in Human Engineering*. New York: The Macmillan Company, 1932.

Starch, D., *How to Develop Your Executive Ability*. New York: Harper & Brothers, 1943.

———, et al., *Controlling Human Behavior*. New York: The Macmillan Company, 1937.

Webb, E. T., and J. J. B. Morgan, *Strategy in Handling People*. Garden City, N.Y.: Garden City Publishing Co., Inc., 1932.

White, W., *The Psychology of Dealing with People*. New York: The Macmillan Company, 1936.

Williams, Whiting, *Mainsprings of Men*. New York: Charles Scribner's Sons, 1925.

42. EFFICIENCY IN WORK

REXFORD HERSEY, PH.D.

Associate Professor of Industry, Wharton School, University of Pennsylvania; Consultant in Mental Hygiene, Pennsylvania Railroad; Associate, Rich and Whitman Associates, Industrial Relations Consultants, Baltimore. Formerly: Director of Executive Development and Training, American Airlines, Inc.; Director of Supervisory Selection and Training, The Glenn L. Martin Company; Personnel Counselor, Richmond Harris, Inc.; Chairman, Committees on Fatigue Research and on Psychological Methods, Industrial Relations Association, Philadelphia; Assistant Superintendent of Safety, A.C. & F. Company; Executive Director, European Student Relief for Central Europe; Personnel Advisor, the German State Railways; Chief of Editorial Section, International Labour Office, League of Nations.

Three psychological techniques essentially affect efficiency in work: motion-and-time study, training, and incentives. Others also affect it, but these three are linked in a close fashion. Contrary to European practice, in the United States the psychologist has not as a rule included motion and time study with their accompanying financial incentives in his tool kit as he has training techniques and the nonfinancial incentives. The reason the field of motion and time study is largely preempted by engineers, perhaps, lies in the fact that the pioneer work in time study was done by an engineer, Frederick W. Taylor.

MOTION AND TIME STUDY

When Taylor originated time study in the Midvale Steel Company in 1881 it was mainly used for rate-setting. Motion study was developed later by Frank Gilbreth and was employed largely to improve methods of work. Today, experience and common practice require that both be used together since one supplements the other.

Barnes (1949) defines motion and time study as the analysis of the methods, the materials, the tools, and equipment used in the performance of a piece of work. This analysis is carried on with the purpose of: (1) finding the most economical way of doing the work; (2) standardizing the methods, materials, tools, and equipment; (3) determining the time required by the average worker to do the task; and (4) training the worker in the new method. The first three of these purposes are discussed in this section; the fourth, training, will be the subject of the following section.

ECONOMY OF METHOD

The most economical way of performing an operation is determined by studying the movements of the worker while performing the job. This is called motion study. All unnecessary motions are eliminated and a sequence of movements most useful for maximum efficiency is determined. The elements surrounding the worker, i.e., tools and materials, lighting, vibration, and noise also must be considered. All conditions should be adjusted to give the greatest comfort to the worker and to derive the greatest net savings in effort and movement.

An effort was made in earlier days to find the one best way of doing a job, and the worker, no matter what his size or physical traits, was supposed to do the job most efficiently by adapting himself to that one best way. Today, it is deemed important to consider the method in relation to the operator, certain variations from standards being permitted under different conditions (Farmer, 1926; Barnes, 1949). The main purpose is to achieve uniform difficulty of achievement. Considering the question of standard procedure more fully, one may even conclude that those practices and standards which are best for a large office or department may not be best for the small office or department.

STANDARDIZING AND TIMING THE JOB

The next step is to standardize the best methods and the job conditions. After these standards have been determined, time study of each small element of the job can then be used to determine the standard number of minutes or hours an average worker takes to perform the operation. This step is commonly referred to as rate-setting. The skill and effort exhibited by the operators during the stopwatch study are rated or evaluated by the time-study observer and the actual time is adjusted by means of this rating factor. The purpose of this rating is to make sure that an average operator can easily do the work in the specified time and to overcome limited numbers of observations and workers observed, that is, inadequate sampling. This corrected time is called base time. To base time are added allowances for personal time, fatigue and delay, the result being the standard time for the task. Today, the standard time (Barnes, 1947), is usually set at such a level that from 25 to 50 per cent of the working force can readily do a third more work than the standard required. Every first-class employee, therefore, can do more work than the standard demands and thus earn extra wages.

ILLUSTRATIVE PROCEDURES

The methods and results of motion and time studies are illustrated in the work of Taylor and the Gilbreths.

TAYLORISM. Taylor (1898) undertook to improve methods in various parts of the Bethlehem Steel Works. About 500 men were chiefly engaged in shovelling. Each good shoveler owned his own shovel. Taylor found that shovelers were lifting loads of about 3½ lbs. when handling rice coal and up to 38 lbs. when moving ore. He found that shovel loads of 21½ lbs. enabled a first-class shoveler to move the most material in a day. By inaugurating a planning department to arrange the work in advance, dividing the men into small groups and providing the appropriate sizes of shovels, each man could work on the type of material with the type of shovel that best suited that material.

The amount of material handled by each man was measured or weighed at the end of the day and a bonus was paid when he did a specified amount of work. If a man failed to earn a bonus, an instructor showed him how to do his job in such a way that he could earn a bonus. In 3½ years, this system reduced the cost of handling material from between 7 and 8 cents to between 3 and 4 cents per ton and the number of men required from around 500 to 140. In spite of all the added staff work and the clerical work involved in the system, a saving of $39,000 was shown during the last 6-month period.

In another experiment Taylor studied the handling of pig iron. The same scientific procedure was pursued, but he went even further and showed that he had an appreciation of the human factor involved. He took into consideration the fatiguing nature of the work and introduced rest periods to prevent it, and to keep his workers producing at a high pitch. It is unfortunate that his successors did not have the broad vision and the truly scientific mind that Taylor possessed. They neglected the human element. They saw in the stop watch only a timing device, only a rigid method of preventing workers from soldiering on the job, and thus brought on merited opposition on the part of labor.

GILBRETHISM. A notable advance in the science of work was provided by Frank Gilbreth (1917) in the building trades. He noticed that craftsmen, particularly the bricklayers, did not work alike. Furthermore, a bricklayer might use different sets of motions at different speeds. Gilbreth took photographs of bricklayers at work and from a study of these photographs was able to make improvements in methods. He invented a scaffold which could be easily and quickly raised or lowered, thus permitting the bricklayer to work at the most convenient level at all times. On this scaffold he placed a shelf to hold the brick and mortar at the most convenient heights. Through transportation and inspection of the bricks by low-grade workers, through their most convenient arrangement on the scaffold, through determining the proper consistency for the mortar and thus eliminating the motion of tapping the bricks into place by the trowel, he was able to increase the number of bricks laid per man per hour from 120 to 350 bricks.

MICRO-MOTION TECHNIQUES

When Frank and Lillian Gilbreth started using motion pictures for the study of skills, their work was along original lines. Micromotion study is the study of the fundamental elements or subdivisions of an operation by means of a motion picture camera and a timing device. This technique was first introduced by Frank Gilbreth at a meeting of the Society of Mechanical Engineers of 1912. It makes possible an analysis of the elementary motions as recorded on the film and the assignment of a time value to each.

CHRONOCYCLOGRAPH. The Gilbreths later developed another method of motion analysis which was called the chronocyclograph. By attaching a small electric light bulb to a finger, hand, or other part of the body, and photographing with a stereoscopic camera the path of light as it moved through space, the path of motion of the operator is recorded in three dimensions. Such a record is called a cyclograph. If the circuit is broken at regular intervals, the spots of light will be spaced according to the speed of the movement. Time, speed, acceleration, retardation, and the direction and path of the motion may be recorded from this graph. This record is called a chronocyclograph. From it, for training or other purposes, accurate wire models of the motion parts are constructed.

THERBLIGS. The greatest psychological contribution of the Gilbreths was their concept of basic divisions of accomplishment. Each division was called a therblig, which is Gilbreth spelled backwards. These therbligs are the basic operations by means of which, in varying sequences and combinations, all operations are performed. The Gilbreths described therbligs. Other investigators have used slightly different subdivisions of work but none have changed the basic conception. Lowry, Maynard, and Stegemerten (1940) used the list reproduced in Table 1.

Effective motion study is impossible without a breakdown of the job into its subdivisions to enable inefficiency to be detected and eliminated. The therbligs of change direction, hold, pre-position, position, search, select, and the delays are all recognized as time consuming and should be eliminated or reduced to the minimum.

Table 1: Basic Divisions of Accomplishment

Purely Physical		Semimental	
Transport empty	TE	Position	P
Transport loaded	TL	Search	S
Change direction	CD	Select	SE
Grasp	G	Objective Basic	
Hold	H	Use	U
Release	R	Assemble	A
Preposition	PP	Delay Basic	
Mental Basic		Avoidable delay	AD
Plan	PL	Unavoidable delay	UD
Inspect	I	Rest to overcome fatigue	F
		Balancing delay	BD

LAWS OF MOTION ECONOMY

The next step in motion study is to apply the laws of motion economy to the building up of the best possible methods practical at the time. These (Barnes, 1949) may be divided in three subdivisions: (1) principles of motion economy as related to the use of the human body; (2) principles of motion economy as related to the arrangement of the work place; and (3) principles of motion economy as related to the design of tools and equipment.

The principles related to the use of the body are of chief interest here: (1) the two hands should begin as well as complete their therbligs at the same instant; (2) the two hands should not be idle at the same instant except during rest period; (3) motions of the arms should be made in opposite and symmetrical directions, instead of in the same direction, and should be made simultaneously; (4) hand motions should be confined to the lowest classification with which it is possible to perform the work satisfactorily; (5) momentum should be employed to assist the worker wherever possible, and it should be reduced to a minimum if it must be overcome by muscular effort; (6) continuous curved motions are preferable to straight-line motions involving sudden and sharp changes in direction; (7) ballistic movements are faster, easier, and more accurate than restricted, fixation, or controlled movements; and (8) rhythm is essential to the smooth and automatic performance of an operation and the work should be arranged to permit an easy and natural rhythm wherever possible.

As an accompaniment to this part of his work the motion study man studies the characteristics of motion. A useful, though by no means complete, classification of movements (Lowry, Maynard, and Stegemerten, 1940) follows: (1) finger motions; (2) finger and wrist motions; (3) finger, wrist, and forearm motions where (*a*) elbow acts as a pivot, (*b*) elbow acts as a ball and socket joint, and (*c*) there are torsional movements; (4) finger, wrist, forearm, and upperarm motions where (*a*) shoulder acts as a pivot and (*b*) there are torsional movements; (5) finger, wrist, forearm, upper-arm, and body motions where there are (*a*) trunk motions on hip pivot, (*b*) trunk motions on ankle pivot, (*c*) leg motions—(c_1) forward or backward and (c_2) side—(*d*) knee motions, (*e*) ankle motions, and (*f*) combination motions.

APPLICATION AND CRITICISM

Applications and adaptations of the principles of motion and time study are being made in many fields, but they must be combined to be successful with the careful analysis of the entire manufacturing process. Their practical value is clearly recognized in spite of many objections that have been raised against them. The chief objections are: (1) the analysis of the parts of a working process does not constitute an analysis of the total operation, i.e., the best complete movement for an operation may not be a combination of elements selected from different workers on the basis of speed; (2) motion and time study as a method of job analysis ignores an important aspect of the job situation, namely, the social setting in which it is placed; and (3) it is usually undertaken to increase production, which is the wrong approach, and its sole object should be to reduce fatigue and increase the ease with which an operation can be performed, thereby, if other things then are equal, the operators will set a satisfactory speed.

Practical experience with motion and time study indicates that these objections vary in importance. It is true in theory that the sum of the parts of an operation does not constitute the total operation, but in practice the new combined movement usually takes less time than the one it displaced. The second objection is largely theoretical but the last carries much weight. When motion and time study is imperfectly carried out or where the reasons for it are improperly understood by the workers, it has often retarded natural increases in the rate of production, more than offsetting the original increase it brought about. If rate standards must be set, the cooperation and support of both the workmen and their leaders are necessary. The attainment of cooperation requires a different approach on the part of management than is customary, leading to the reeducation of both workers and supervisors. Whatever its faults, most of which have been based on a misconception of motion and time study as it is today or are directed at the misuse which misguided or unintelligent employers have made of it, the industrial psychologist in the United States can use the technique of motion and time study and its principles to greater advantage than he has to date.

TRAINING

The application of psychological knowledge of the training problem to efficiency in work concerns (1) the assembling of the general principles of training and (2) an understanding of the techniques of their use in maintaining efficiency of workers while on the job, i.e., in the training of employed workers on changing jobs instead of the training of new employees, though the basic techniques are the same.

THE FOREMAN, AGAIN A TEACHER

The early foremen were responsible for the education and training of employees as well as for production. From 1890 to about 1920, they were relieved largely of this responsibility through the establishment of vestibule schools. Since 1920, there has been a distinct swing back to the supervisor as the trainer. He is coached by a specialized industrial educator or the motion study man because though proficient in the knowledge of the work he is not able to impart that knowledge to his men. He must be shown both how to collect and select the training material and how to get it across with a punch. A primary prerequisite which is neglected by the majority of companies is to ground the foremen or workmen-teachers in the principles of motion study and motion economy. This step should lead naturally into worker learning and to applying certain fundamental principles of motor learning.

PRINCIPLES OF MOTOR LEARNING

The following eight principles are important in the training program; the first five of them are adapted from Moore (1939). (1) The best way to exercise a skill must be discovered before it is taught. (2) The job should be taught in psychological units involving easily acquired sequences of movements but the significance of the separate units will be better appreciated if the learner is given a birds-eye view of the new or revamped job. (3) Visual and other mechanical aids to learning should be used when possible. This method facilitates the presentation of an operation and drives home the correct way. (4) Proper incentives must be provided. (5) The initial stages of any new motor habit are extensions of old and similar habits. An examination of all parts of the task familiar to each learner should be the starting point of the retraining process but the most effective training takes place on the actual operations to be performed. Transfer of motor skill seems no more possible than transfer of training in other lines. (6) The laws of habit formation must be observed: (*a*) the learner must be convinced that the task is worthwhile and must be mastered in a certain way; (*b*) the correct start is very important; (*c*) the process of acquiring the habit should be made by developing coarse, large movements first and finer ones later, and (*d*) the motions must be repeated until the correct form is automatic and natural. The training trials, however, must be properly spaced. Hovland (1941) insists that, when there is great pressure for a speed-up of training, this principle must not be neglected, since too rapid skimming over material may actually prolong the period required for training rather than lessen it. The more rapid the operation, the more important it is to distribute the practice. (7) The job must be analyzed to determine whether the performance at slow speed is the same as that at high speed (Hovland, 1941). If it is, accuracy should be stressed; if not, the form of the finally correct performance should be the goal, even if a degree of accuracy is sacrificed, since the correct pattern should be practiced from the start. (8) The foreman or instructor must use a check to evaluate the rate of learning. He can thus encourage the learner by showing him his progress, and can also avoid wasting time overlearning. The application of these principles will enable the foreman, director of training, and the time and motion study man to develop sound content and to work up effective media and techniques of teaching the changed methods and operations.

CURRENT STATUS OF TRAINING PROGRAMS

The status of training in industry today is uncertain. Some companies are adding training departments; others are revamping present departments in line with changed requirements; still others are curtailing training programs. The last survey available (National Industrial Conference Board, 1943) indicated that out of 230 companies covered by the survey, 26.5 per cent conducted training courses, but the survey did not indicate how many of these were primarily for the purpose of training present employees in new or changed operations. An earlier survey of the National Industrial Conference Board (1937) seems to present a picture of training conditions fairly typical of normal years. That survey covered 473 companies with 626,668 employees: (1) four out of five of the companies had a definite training policy; (2) training on the job was given in 98.5 per cent of the companies and only 8.5 per cent had vestibule schools compared with 26.5 per cent in 1943; (3) training was given to mature as well as to young learners; (4) compensation of persons in training was at least 50 per cent of the regular rate in 91 per cent of the companies and 80 per cent of the regular rate or more in about 32 per cent of the companies; (5) the time required for training ranged from one week to more than five years but more than 60 per cent of the companies were able to give any necessary instruction in 6 months or less; (6) definite policies for training employees for versatility were in effect in 47 per cent of the companies; (7) regular systematic apprentice-training was reported by 272, or 57.5 per cent of the companies covered; (8) a large proportion of these plans were in metal-working industries; (9) the number of apprentices in training in these companies was 7,322, or 1.1 per cent of the aggregate employment of the companies. [If all companies in the United States maintained this proportion, 580,000 apprentices should be in training. The Department of Labor, however, estimates that only 140,000 apprentices are in training under organized programs.]; (10) about 55 per cent of the companies reported that apprentice-training provided all the instruction given and about 45 per cent utilized outside facilities, usually for the schoolroom instruction; (11) more than a third of the companies made no definite provision for schoolroom instruction; (12) 4 years was the most common length of apprentice-training courses.

NEED FOR CONTINUOUS TRAINING

The above survey indicates how fully proper and continuous employee training was recognized even before World War II. As a rule a worker does not perform an identical routine operation over a long period. He usually performs many different operations in the course of a week or two. Likewise, constant changes in method and the rapid introduction of new models give to the worker new jobs which he must be continually learning. On simple jobs where the worker needs new instructions for the performance of particular operations, the simple instruction sheet is often satisfactory. The best example of this is the Job Breakdown Sheet for Training Man on New Job developed by TWI (Training Within Industry) during World War II. On it were listed on one half of the page the important steps in the operation. On the other, opposite to the operation, were listed the key points, anything that might make or break the job, injure the worker, make the operation easier to learn or do. Where the work is somewhat more involved, more detailed instruction sheets are prepared, and Williams (1936) used still pictures very successfully in connection with written instructions in an Eastern rubber company.

SAMPLE TRAINING PROGRAM

For those operations which require considerable skill on the part of the operators, the principles of training mentioned above must be more carefully followed whether the operators are trained on the job or in a vestibule school. An excellent example of the possibilities of this type of training is given by the General Electric Company, Erie Works, (1940) in their individual training of machine operators based on motion study, using films.

THE PROBLEM. Owing to the increases in production demand, changes in the manufacturing process, and normal labor turnover, the General Electric Company found that it was constantly necessary to train operators in types of work that were new to them. The need was to find a procedure that would help an inexperienced man or woman to develop the required speed of production in as short a time as possible. At the same time it was hoped that by improving the work methods of experienced operators it would be possible to increase productivity and to decrease individual fatigue.

THE SOLUTION. The problem was approached from two angles: (1) the basic training would have to be done with the individual workers but (2) it was necessary first to give the organization as a whole a consciousness of the value of improved methods. The logical starting point was through motion analysis and a program was designed to make the technical staff and supervisory personnel more motion minded.

During one year, approximately 300 foremen, time study men, tool planners, and tool designers were given the opportunity to study a basic course in motion study. The original group to receive the instruction was made up of one man from each of the company's units. The unit met at the West Lynn plant for intensive training. This central group continues to meet once a year to exchange information on new developments. The instruction then was given at the Erie Works to some fifteen groups that ranged in size from fifteen to twenty workers for 20 weekly 2-hour sessions.

The technique of presentation was a combination of lectures to explain the basic theory, discussion to clarify obscure points, and work projects to give practice in the application of the basic principles. The understanding of the motion study methods gained throughout the organization created a different attitude on the part of both the supervisors and the employees toward improvement techniques. The foremen were able, not only to decrease the initial training time, but also to help veteran workers accomplish more by easier methods.

At the same time that this educational work in the organization as a whole was being carried on, attention was concentrated on the development of skill in particular types of work. In one department, a room was set aside for training to meet a heavy demand for skilled workers in taping of leads and field coils. Veteran workers were given instructor training, special work benches were designed, and camera equipment was made available for film instruction. A careful analysis of the job was made in order to determine the most effective method of performing the various operations. Camera studies were made of the best operators and these pictures were analyzed by foremen and methods men to eliminate waste motions. Once the best method of job performance had been defined, a training procedure was developed to meet the particular job needs. Instruction Sheets describing the motion element and the key points were written and were appended to a series of photographs showing the operation sequence.

The trainees were given an outline of motion economy principles. The instructor performed the first element of the operation slowly until the learner was familiar with the motion sequence. The trainee then took his place at the machine and practised the first element while the instructor pointed out the correct methods. Not until the learner had mastered the motion sequence of the first element was he permitted to attempt the second. This plan was carried out until efficiency was gained in the entire operation. It was found that speed came as a natural result of correct habit formation.

Daily checks were plotted on progress sheets and the results then figured on weekly sheets. The trainees were shown their daily and weekly averages so that they might have objective proof of their rate of improvement. Individual time studies were made and a high rate of learning was apparent in the learner's progress. The following three trainees are reported as typical. One girl had already worked on the job 23 weeks and her rate of production had exceeded the learner's guaranteed rate in only two of the weeks. Five weeks of training increased her efficiency so that she was receiving the anticipated earned rate. Two girls with 5 weeks' shop experience had low averages per hour. After 3 weeks' training they had increased their efficiency to such an extent as to be receiving twice their original rate per hour. Initial training time of new operators was found to have been cut in half by the new methods.

It was reported that improved techniques resulted from this program which decreased job training time to such an extent as to give marked savings in training expenses. The effect of the job instruction in producing an operator who could approach the job with more understanding was seen in a decrease of supervision time required by the operators after training. More complete records could be given the employment office to be used as a basis for rehiring. The training methods offered possibilities for the development of selection tests for hiring new men. A better employee morale resulted from increased individual earnings, the realization by the employees that the company was concerned with individual progress, and the recognition by the operators that the new methods reduced individual fatigue. This example, typical of many which helped later to build up the wartime efficiency of American industry, aptly illustrates what can be accomplished when the proper principles of learning and motion economy are applied to a practical job of training.

INCENTIVES

How to get people to want to work has always been a problem. Four factors are concerned in this problem: (1) the energy available, (2) the goal or goals, (3) the environmental conditions, and (4) the incentives used.

COMPLEXITY OF INCENTIVES

Hovland (1941) points out that numerous companies have found periodic increases in wages during learning effective in increasing speed of learning, but observation of one's own progress, an encouraging word, personal interest shown by the instructor, competition with other learners, realization of the significance of one's work in the total finished product, and patriotism may all be powerful motives. Competition can be skillfully employed in motivating the learners. Continually unsuccessful competition, however, may cause them to become discouraged. It is for this reason that competition against one's own record often is a less dangerous incentive. Hay of the Pennsylvania Company (1941) combined the incentive of showing workers their own results with one of promotion and produced very striking improvement.

Planty (1946) insists that the desire to learn a particular job is increased, not only by the incentives referred to above by Hovland, but also by encouraging the employees to their fullest self-development as men and women. An example was a plant where all woman employees were being given training in industrial hygiene, clothes-design and selection (taught by specialist from Lord & Taylor), makeup (taught by Richard Hudnut), posture, business etiquette, etc. The example carries the principle of self-development to somewhat extravagant limits but basically it is sound and must be considered as part of training incentives.

At any one time many incentives, both from within and without, motivate the worker, influence his general emotional state and his will and capacity to work. Not only do different incentives appeal to different workers, but the same worker may be receptive to different incentives at different times during the same day. The most common incentive is money.

WAGE PAYMENT SYSTEMS

Many different wage incentive plans are being used. The straight piece rate, which is payment at a fixed rate per unit produced, is the simplest and oldest. Its modifications are divided into three groups (Alford, 1942): (1) the Taylor plan pays a lower rate per piece to each worker who does less than a fixed amount and a higher rate for more than that amount; (2) the task and bonus rate, which was developed to avoid the penalty imposed on workers who fall below the standard amount, pays a guaranteed minimum for the day and if production equals or exceeds the standard he is paid at a piece rate; and (3) another form of the task and bonus rate is based upon two variables, usually a ratio between a standard rate and an acquired production rate. The plans of Taylor, Merrick, and Gantt are included in the first group; those of Halsey, Diemer, Baum, Bedeaux, Sherman, Rowan and Barth, in the second group; and those of Emerson, Wunnerlund, Knoeppel, Bigelow and Parkhurst, in the third group. All of these plans have been used successfully. A survey (1940) covering 313 companies showed that about 60 per cent of the incentive workers were paid by some form of piece rate or tonnage rate system, about 31 per cent under a premium or bonus plan, and about 9 per cent under measured day work.

Under the system of measured day-work an hourly rate is set up, which reflects in part the requirements of the job and in part the qualities of the worker, such as productivity, quantity, dependability, and versatility.

EMPLOYEE ATTITUDES

These plans depend for their success on the accuracy of the measurements upon which they are based. Unjust wage payment systems have led to much friction and rate cutting and even in the best managed companies workers are usually fearful of them. The feeling of antagonism toward the piece rate system is still marked in many departments of a large corporation where the method of setting the rates is exceedingly accurate and the workers' representatives are members of the rate setting committees. Where the relations with the foreman are not of the best, the workers practise a definite restriction of output so that their earnings will not go above a certain agreed-upon percentage, which averages 20 per cent to 23 per cent above the hourly rate, instead of the 27 per cent to 30 per cent that similar workers in other departments are earning. In spite of the fact that financial incentives have increased the rate of production, decreased costs, and increased the financial return to the workers, they and the motion and time study on which they are based have not brought the worker generally to feel that he can put forth his best production without fear that he will work himself out of a job or suffer some other misfortune instead of being properly rewarded.

A fairly recent National Industrial Conference Board report, however, shows an increase of favorable attitude in approximately 85 per cent of the companies studied toward the techniques of motion and time study; 9 per cent reported a decrease; and 6 per cent no change. The following reasons for the improvement are given: (1) avoidance of complicated systems which cause misunderstandings; (2) improved rate-setting techniques which won the worker's confidence in the ability of trained experts to set fair task standards; (3) selling employees on the value of accurate measurements and the elimination of guess work; and (4) explanation of the procedures involved in order to remove the mystery from motion and time study.

SUCCESS OF AN INCENTIVE WAGE SYSTEM

Smith (1931) cites the following conditions which must be met before financial incentives will produce the best results: (1) they must be related to individual skill and performance; (2) their effectiveness will depend on the faithfulness with which they are followed as a basis for payment, promotion, or discharge; (3) they must represent an appreciable addition to the standard or market wage rate; (4) they must be accompanied by a careful selection, training, and promotion program; (5) the fear of insecurity and of loss of prestige must be removed; (6) they must be simple and easy for the worker to calculate his earnings; and (7) they must provide rewards quickly after they are earned.

Other problems, however, must be solved before an incentive plan will function with complete satisfaction. One of these is that of guaranteeing incentive standards. After the damage to worker attitude caused by the era of rate-cutting it was commonly considered proper practice to guarantee rates for an unlimited period. Today, many firms are adopting a limited guarantee, the rates being subject to review upon the expiration of perhaps a year or upon a change of method, whichever occurs first. Another problem to be solved is rating the worker's physical skills and efforts. It is desirable that the workers keep their top performance a reasonable distance below a killing pace, yet do a "standard day's work." Two interesting possibilities of solving this problem, exclusive of rating, are in process of development: (1) the standard data method, which relies on thousands of time studies of common elements that can be assembled as a basis for determining the correct time allowance for new operators, and (2) methods-time measurement, which serves the same purpose (Maynard, Stegemerten, and Schwab, 1948) but relies on the motion-picture camera and 16 mm. film. It is yet too soon to predict the success of these methods.

PREVAILING PRACTICES

The National Industrial Conference Board makes the following report on prevailing practices in the administration of wage incentives: (1) replies received from 293 companies indicates that in 186 companies, or 63.5 per cent, the standards are definitely guaranteed for an unlimited period of time so long as a method change does not automatically require cancellation; in twenty companies, or 6.8 per cent, the policy is to guarantee standards for a limited period of time, i.e., in seventeen companies one year, in two companies 3 years, and in one company 6 months; (2) of two hundred and sixty nine companies that reported their policy regarding the payment of workers of make-up allowance for delays that are not the fault of the worker, about 78 per cent calculate the make-up payments on the basis of individual jobs and 22 per cent take into consideration average earnings for the week in calculating the make-up allowance; (3) in seventeen companies, or about one eighth of the unionized companies, the union participates in the establishment of wage incentive standards; (4) 75 per cent of 900 manufacturing companies having from 100 to 10,000 employees each and being engaged in a type of manufacturing which offered an even choice of either day work or incentive systems, have some form of wage incentive plan. These companies were selected from 2,700 companies in all kinds of business from a recent Conference Board survey which showed for the whole only 51.7 per cent using incentive plans.

DEVELOPMENTS

Two directly opposed methods of wage payment have existed for centuries: ordinary time wages and ordinary piece wages. One took no account of output and the other took no account of time. There is a noticeable trend, today, back to those two simple systems with the important difference that the gap between them has been narrowed. Outstanding developments during the past forty-five years have been in devising tools to aid in the administration of wage incentive plans: time study, motion study, and appraising the worker's performance, rather than in any particular wage incentive plan. Hand in hand with this has gone the professional development of skill in the use of these tools. The tools of measurement that have been developed for incentive plans have been found to be almost equally applicable to both piece wages and time wages and these tools have taken an important place in the industrial world.

OTHER INCENTIVES

Other financial incentives than those based upon the immediate return from production that have been tried include suggestion awards, quality bonuses, stock purchase plans, savings plans, group bonuses, and profit sharing. Blanket criticism of these incentives is as unwarranted as blanket praise. All may be helpful depending on the situation and the phase of the business cycle. On large jobs, where from 6 to 12 men work together under the direction of a leader and the various tasks are apportioned among the various members, group incentives have proved equally as successful as the individual financial incentives previously described, provided adequate safeguards are maintained. The quality of leadership displayed by supervisory personnel cannot be minimized in the evaluation of final group output.

Financial incentives are but one of the motivating forces of mankind through an important one to civilized man. A number of incentives influence the worker's progress in training. The worker is surrounded by many physical and ego desires which influence his will to work and readiness to put forth energy in accomplishment. Various incentives play upon the worker at all times, influencing him through his emotional attitudes, his interest, and his morale.

POINT OF VIEW

The level and consistency of production are determined by a wide variety of variables which require careful study and control for the attainment of maximum yield, satisfaction, and fitness on the part of the individual worker. The entire industrial environment of the worker, as well as his physical condition and outside relationships, are important determiners of his efficiency. The exact value of each of these factors and the pattern of their interplay are questions which are far from fully answered. The complete answer may never be given, but much progress is being made and will continue to be made in understanding efficiency in work.

READING REFERENCES

Alford, L. P., and J. R. Bangs, *Production Handbook*. New York: The Ronald Press Company, 1944.

American Management Association, "Principles and Methods of Industrial Training." *Personnel Series*, 1941, No. 47, 1–14; "Compensation Plans for Executives and Workers." *Personnel Series*, 1940, No. 30, 17–28; "The Future of Industrial Training." *Production Series*, 1946, No. 167, 25–34.

Balderston, C. C.; V. S. Karabasz; and R. P. Brecht, *Management of an Enterprise* (3rd ed.). New York: Prentice-Hall, Inc., 1949, pp. 115–165, 397–425.

Barnes, R. M., *Motion and Time Study* (3rd ed.). New York: John Wiley & Sons, Inc., 1949, pp. 1–404.

Gilbreth, F. B., and Lillian Gilbreth, *Applied Motion Study*. New York: The Macmillan Company, 1919, pp. 42–50.

Halsey, G. D., *Training Employees*. New York: Harper & Brothers, 1949, pp. 123–148.

Lowry, S. M.; H. B. Maynard; and G. J. Stegemerten, *Time and Motion Study and Formulas for Wage Incentives*. New York: McGraw-Hill Book Company, Inc., 1940, pp. 74–136.

Maynard, H. B., and G. J. Stegemerten, *Operation Analysis*. New York: McGraw-Hill Book Company, Inc., 1939, pp. 213–228.

Maynard, H. B.; G. J. Stegemerten; and J. L. Schwab, *Methods-Time Measurement*. New York: McGraw-Hill Book Company, Inc., 1948, pp. 1–41.

National Industrial Conference Board, *Personnel Practices in Factory and Office*. New York: The Board, Studies in Personnel Policy, No. 23, 1940; No. 59, 1943; No. 88, 1948.

———, *Quick Training Procedures*. New York: The Board, Studies in Personnel Policy, No. 26, 1940.

———, *Selecting, Training, and Upgrading Supervisors, Instructors, Production Workers*. New York: The Board, Studies in Personnel Policy, No. 37, 1941.

———, *Some Problems in Wage Incentive Administration*. New York: The Board, Studies in Personnel Policy, No. 19, 1940.

———, *Training for Industry*. New York: The Board, Research Report, No. 37, August, 1937.

Planty, Earl G.; W. S. McCord; and C. A. Efferson, *Training Employees and Managers*. New York: The Ronald Press Company, 1948, pp. 29–109, 200–213.

43. OCCUPATIONAL SAFETY

GILBERT E. TEAL

Lt. Colonel and Deputy Chief, Ground Safety Division, Hq., US Air Force. Formerly: Supervisor of Motor Vehicle Units with First, Third and Fifth Air Forces; Safety Engineer, Engineering and Inspection Division, Travelers' Insurance Company; Safety Supervisor, Naugatuck Footwear Plant, US Rubber Company.

ANNUAL COST OF ACCIDENTS

Accidents in American Industry annually drain huge segments of profit from the coffers of productive enterprise. The National Safety Council (1947) estimated total accident costs for the year 1946 alone at $6,400,000,000, of which industrial accidents accounted for $2,400,000,000. Accidental deaths for 1946, an average year, totalled 99,000, and 10,400,000 persons were reported injured during the same period, 370,000 of whom received permanent physical impairment. 1946 was just an average peacetime year—with a relatively normal exposure hazard. 1946 loss figures continue to follow closely previous yearly experience trends, in spite of the concentrated efforts of technician and layman to reduce the annual toll.

HISTORICAL BACKGROUND

Control of occupational accidents has received ever-increasing attention during the past forty years. Initial control measures stemmed from the activities of the Association of Iron and Steel Electrical Engineers (Holbrook, 1939) founded in 1907. This organization sponsored a meeting in Milwaukee, Wisconsin, in 1912, the outgrowth of which was the formation in 1913 of the National Council for Industrial Safety, subsequently renamed the National Safety Council. Initial attention was directed entirely toward mechanical safeguards for the prevention of bodily injury to workmen. The advent of World War I, and the urgent necessity for conserving manpower and resources to conduct the war gave added impetus to the growth of the safety movement. Personnel classification techniques contributed measurably to the understanding of safety problems. As early as 1919, studies of accident causes were made in England by Greenwood and Woods and their conclusions were presented before the Royal Statistical Society. Studies of accident causation have been conducted in various countries and reported by Marbe (1923, 1926), Newbold (1926), Schmitt (1926), Hildebrant (1928), Farmer and Chambers (1926), Hewes (1921), Fisher (1922), and many contemporaries during the decade following World War I.

The initial scientific literary presentation devoted entirely to the problem of accident prevention techniques made its appearance in the United States under the title "Industrial Accident Prevention," in 1931. This publication by Heinrich was the first effective delineation of the basic factors of accident frequency, cause, analysis, and control by the safety engineering group, and the first open recognition of the problem of accident control as a basically psychological problem, subject to clinical techniques, psychometric determinations, and established personnel control procedures.

BASIC PHILOSOPHIES OF ACCIDENT PREVENTION

Certain of the basic philosophies expounded by Heinrich (1931) are worthy of further attention. The basic cause of 88 per cent of all accidents resulting in personnel injury were attributed by Heinrich partly or wholly to the unsafe acts of persons. Blake (1944) criticises such an arbitrary determination of percentages as invalid, but concurs in the principle of the preponderance of accidents being attributable to the fault of persons. Studies of aircraft accidents reported by McFarland (1946), the Metropolitan Life Insurance Company (1930, 1931) Bingham (1927), and the Center for Safety Education (1941) substantiate the conclusions cited.

A second basic philosophy established by Heinrich, and later accepted generally by psychologist, economist, cost accountant, and engineer, is the theory of the indirect-direct accident ratio: "The indirect cost to the employer for any accident is four times the total direct cost." This indirect cost ratio is confirmed in later writings by Spriegel and Schulz (1942).

Third among the basic conclusions of Heinrich was the belief that accidents were not distributed on a chance basis among a given population. Bingham's research on the Boston Elevated Railway (1928), and studies completed by the Metropolitan Life Insurance Company were cited by Heinrich in substantiation of his hypothesis that basic causes of accident proneness existed and that further research could detect and correct these causes, utilizing existing clinical techniques. The theory of accident proneness is amply illustrated in the experiments of Marbe (1923) and many others, reports of which are contained in condensed form in the works of Viteles (1932), Tiffin (1942), and Maier (1946).

ESTABLISHED PSYCHOLOGICAL RELATIONSHIPS IN ACCIDENT CONTROL

Certain definite conclusions have been reached by both the psychologist and engineer in the continuing study of accident control techniques. These established, proven relationships may be summarized as follows:

ACCIDENT PRONENESS

Conclusions indicate a definite predeliction of certain individuals toward accidents under uniform conditions of work environment. Bingham's (1927) experiments established the facts that 85 per cent of the employees investigated had an average of 2.3 collisions, while the 15 per cent remaining accounted for an average of 8.2 collisions. Other experiments tended to confirm these facts, and it is now generally accepted that 25 per cent of the employees are responsible for 75 per cent of the accidents occurring in a given exposure. Farmer and Chambers (1929) further conclude that these accident repeaters possess "a poor aestheto-kinetic coordination," but Bingham (1927) claims that this may be compensated for in partial measure through clinical approach and applied corrective educational techniques.

AGE IN RELATION TO ACCIDENT RATE

Tiffin (1947) reports a definite correlation between age and accident rate, summarizing conclusion of Hewes (1921), Gates (1920), Schmitt (1926), and Lipmann (1925). Tiffin (1947) indicates that the number of hospital visits per employee rises sharply to a peak between the ages of eighteen and twenty-three, and then declines gradually to age sixty. Viteles (1932) calls attention to the clear connection between age increment increase and increased length of service (or job experience) and accident incidence, but reports that the relationship between decreased accident rate and increased age remains undisturbed when length of service is held constant. When age is held constant the association between decreased accidents and increased service tends to diminish. Hence, Viteles concludes, the relationship between increased age and decreased accident rate is a stronger tendency than the apparent similar relationship between experience and decreased accident rate. The foregoing results are further substantiated by statistics of the National Safety Council (1947) shown in Table 1.

Table 1: Accidental Deaths by Age, Sex, and Type, 1945

Age and Sex	Accidental Deaths All Causes
All Ages	95,918
Under 5	7,741
5 to 14	6,836
15 to 24	12,446
25 to 34	9,737
35 to 44	9,656
45 to 54	9,661
55 to 64	10,436
65 to 74	11,238
All others	18,167
Sex	
Male	66,339
Female	29,579
Per cent female	31%

EXPERIENCE CORRELATION WITH ACCIDENT RATE

Fisher (1922), Chaney and Hanna (1918), Dussier (1930), and Tiffin (1947) all conclude that a definite relationship exists between increased length of service and decrease in accident rate although the rule is not held invariable. Certain studies by Shrosbree (1933), Stevens (1929), and Newbold (1926) indicate a positive correlation between increased length of service and increased accident rate. Statistics quoted in Table 1 show a positive acceleration in accident rate among persons over forty-five, which would superficially tend to substantiate the latter findings. Further analysis of the figures in Table 1, however, indicate the increased accident rate above 45 is not attributable to industrial accidents, but to such increased types as accidental falls, burns, and similar off-job accidents. Table 2 from Fisher (1922) indicates a typical study substantiating the conclusions that increased experience results in a decreased accident rate. The critical period of one month indicated in Table 2, and other similar analyses lead Fisher to conclude that employees with less than one month's service are six times as liable to accidental injury as those with more than one month's service.

Table 2: Rate of Accidents in Proportion to the Average Number of Employees in Each Length of Service Group (Fisher, 1922)

Length of Service	Number of Employees	Number of Accidents	Accident Rate
Less than 1 month	41	74	181%
1 to less than 3 months	96	121	127
3 to less than 8 months	249	217	87
8 to less than 12 months	183	114	62
Total less than 1 year	569	526	92
1 to under 5 years	750	430	57
5 to under 10 years	267	125	47
10 to under 20 years	223	119	53
20 to under 30 years	164	59	36
30 to under 40 years	75	33	44
40 to under 50 years	33	22	67
50 to under 60 years	4	0	0

Various explanations have been offered for the phenomena cited, such as elimination of the reckless employee early in his career, progressive advancement to safer jobs, increased economic and family responsibility, and numerous other conclusions. The preponderant weight of data indicates, however, a definite positive correlation between increased service and decreased accident rate.

MENTAL ABILITY AND ACCIDENT RATE

Recognition of hazards inherent in the performance of various jobs indicates that level of general ability is important in accident causation. But it would be impossible from this point of view to ascribe specific standards or levels of achievement, mental age, or intelligence as a prerequisite for industrial employment. Brody (1941) aptly summarizes the situation in the conclusions of his research study in the following manner: "Obviously an idiot cannot be trusted with a car. On the other hand, a high degree of intelligence does not seem to be essential to safe and skillful operation of a motor vehicle." Henig (1926) reports a definite correlation between accidents and Army Alpha test scores. Snow (1926) indicates combined intelligence and emotional stability scores were more predictive than emotional stability tests alone. Maier (1946) concludes from a study by Selling (1943) that when available intelligence level is inadequate for the situation the trait of accident proneness is present.

PHYSICAL CONDITION AND ACCIDENT RATE

A susceptibility to accident may be directly connected with a generally lowered physiological effectiveness, according to Viteles (1932). Newbold (1926) substantiates the conclusion that a direct relationship exists between minor illnesses and accident rate. There is a suggestion that such correlation may be connected with emotional instability or hypocondria, which in turn are correlated, according to Brody (1941) and Maier (1946), with accident rate.

Much attention has been devoted during recent years to the subject of "physically impaired" employees, in contrast to "physically unfit" employees discussed in the preceding paragraph. Studies by the National Safety Council (1947) indicate that physically impaired workers are no worse accident risks in manufacturing industries than workers without impairments performing the same work. Properly placed, with emphasis on the work capabilities of the impaired worker, physical impairment is not an accident factor. With the increasing search for employment on the part of disabled veterans of the recent war, careful attention should be given to the policies of proper selection techniques and placement as a therapeutic device. Personnel, properly placed with a scientific diagnostic employment procedure, are a valuable source of increased industrial manpower. The Association of Casualty and Surety Executives (1945) and Bridges (1946) have recently published excellent guide books on the subject of job placement of the physically handicapped-physically fit personnel.

SEX AND ACCIDENT RATE

It is generally concluded that women are more inclined to have industrial accidents than are men. Such is the conclusion of Viteles and Gardner (1929), Maier (1946), and other writers. Such items as periodic physical depression, increased susceptibility to fatigue and the effects of temperature are thought to make it difficult for female employees to maintain accident rates as low as those of male workers. These and other deficiencies in strength, mechanical adaptation, and similar hindrances, however, are normally compensated for by selection and placement of female personnel on less hazardous jobs.

An unpublished study by Teal of the accident experience, frequency, and severity among female labor compared with male labor shows relative hazard ratios to be more favorable to female workers than generally concluded. Results of the analysis, covering five years of experience in a large rubber footwear manufacturing plant, for which the average plant population was 2,044 male and 2,623 female workers, indicates that the sex of workers in the rubber industry has little or no effect on the relative susceptibility of the individual to minor accidents. Of 7,005 accidents occurring during 1939, 44 per cent of the total employees (male) accounted for 42 per cent of the accidents while the remaining 56 per cent of the plant employees (female) accounted for the remaining 58 per cent of the total injuries. A five-year survey of lost time accidents (1935–39) with a similar population breakdown showed the 56 per cent female operators having only 26 per cent of the lost time accidents while the 44 per cent male population accounted for the remaining 74 per cent. A four-year study (1936–39) of direct accident costs shows an even greater differentiation. The 44 per cent male employee population accounted for 91.5 per cent and the 56 per cent female population the remaining 8.5 per cent of the days lost in the same four-year period. Cases of permanent partial disability were apportioned 93.7 per cent to male employees and 6.3 per cent to female. Statistics cited in the foregoing are shown in Table 3.

SENSORY MOTOR COORDINATION

Much stress has been placed upon the connection between accident rate and speed of reaction of individuals. Many test devices have been evolved by technician and layman alike for the purpose of measuring such time elements as speed of application of an automobile brake, stopping time for the arresting of a freely falling object, dotting tests, and similar devices. The Aetna Life Insurance Company, the American Automobile Association (1938) and others have evolved whole batteries of mechanical predictive devices intended to discriminate between good and poor motor vehicle operators. Unquestionably many of these devices are partly predictive. Certainly there appears to exist a definite relationship between such items as poor coordination, poor vision, poor reaction time, and the parallel poor driving experience. It is becoming increasingly evident, however, that sensory motor tests, to be effective, must measure not only speed, but correctness of reaction. Tremendous strides in the direction of sensory motor test techniques were made in connection with the selection of pilots and air crew personnel in the military and naval Air Forces during World War II. Research conducted by the Army's Aero Medical Laboratory and School of Aviation Medicine (Baitsell, 1946) was so effective that the test battery administered could predict with considerable accuracy the probability of the applicant successfully completing the pilot training. Applicants were separated into nine stanine score groups and those in the upper group had only 6.4 per cent washouts while those in the lowest group had a washout rate of 81.5 per cent. (See Figure 1.)

EMOTIONAL ADJUSTMENT

Proper placement is essential in any job. The problem of the employer is to determine the characteristics of the employee in order to properly utilize his services. Existing studies by Viteles (1932), Tiffin (1942), and Maier (1946) indicate clearly the correlation between intelligence level and ability to withstand monotony and escape boredom. The importance of attitudes in connection with successful accomplishment of an employee's job has been discussed by Bingham (1927) and numerous con-

Table 3: Comparison of Male and Female Accident Experience

	1939		1938		1937		1936		1935	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Lost time accidents	7	5	7	1	11	2	11	5	18	6
Days lost	5,063	106	4,629	35	1,281	86	537	549	Undetermined	
Direct cost of l.t. accidents	$8215.48	341.12	7817.34	190.00	2005.46	237.38	1279.78	1009.69	Undetermined	
Number of minor accidents	7,005		5,304		8,857		7,033		6,868	

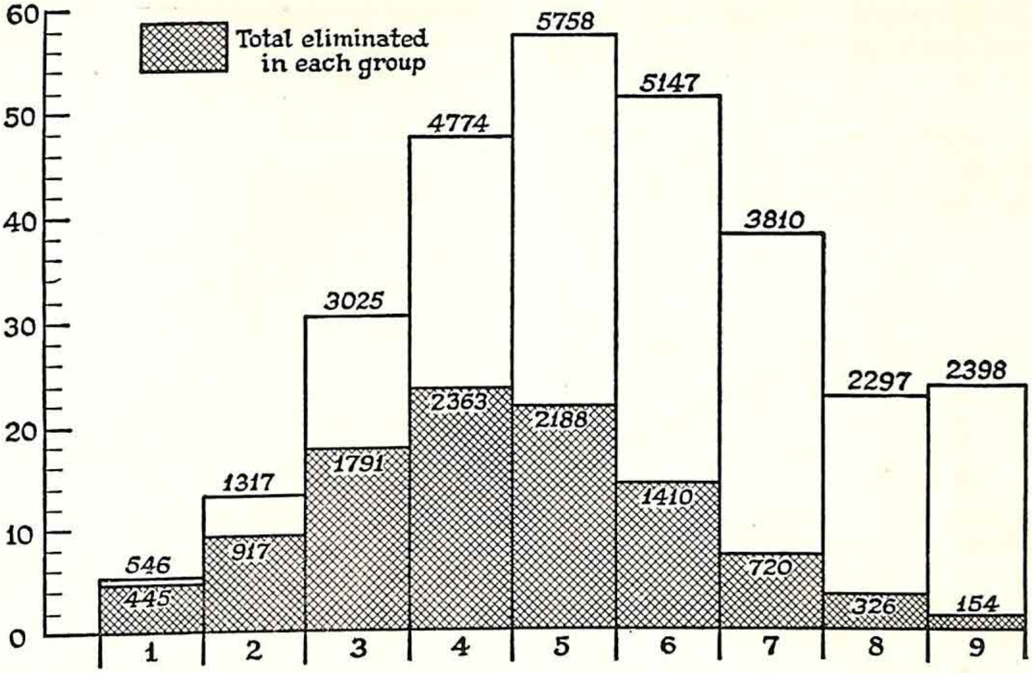


Figure 1 (VI–43): *Prediction of Pilot Success by Total Test Battery*

temporaries. Emotional maladjustment is a contributing factor in untold numbers of accidents. Brody (1941) reports maladjustment in general is much more common among accident repeaters. Utilizing the areas of adjustment of the Bell Adjustment Inventory—adult form—Brody obtained the results shown in Figure 2.

BLOOD PRESSURE AND ACCIDENTS

One study by Brody indicates a definite correlation between low systolic blood pressure and accident proneness. Brody concludes that other factors, such as age, experience, and training, being equal, lowness of systolic blood pressure may be symptomatic of accident proneness. The corrollary symptoms—usually mental and physical exhaustion—are of such a nature as to infer the validity of Brody's conclusion. Bridges (1946) indicates such low blood pressure may reduce capability for work in some people, but that it is not an important industrial problem. With respect to high blood pressure, ranging between 170 and 200, Bridges (1946) recommends

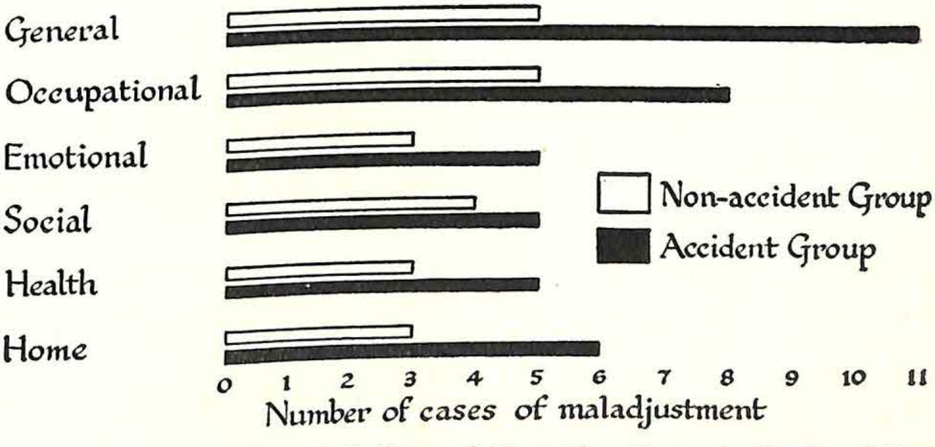


Figure 2 (VI–43): *Comparison of Accident and Nonaccident Groups in Number of Cases of Maladjustment on the Bell Adjustment Inventory* (from Brody, 1941)

avoidance of jobs requiring work at heights, stooping, kneeling, jumping, climbing, and heavy lifting. Where blood pressure exceeds 200, all jobs are considered hazardous except for carefully selected sedentary occupations. Detection of high or low blood pressure is a simple matter to determine medically, and placement should be made according to requirements once such physical condition has been discovered.

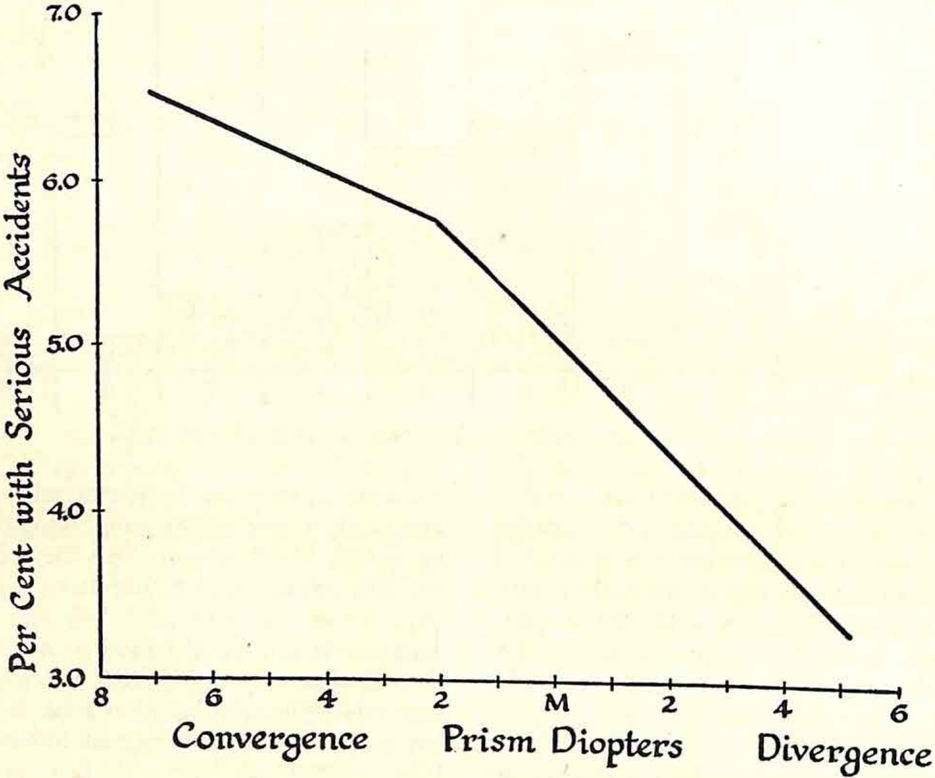


Figure 3 (VI–43): *Accident Proneness and Vision, from a Chart by Tiffin* (1947, p. 209)

VISION AND ACCIDENT RATE

The subject of vision, considering the all-inclusive term to encompass visual acuity, peripheral vision, color blindness and all of the ramification of visual measurement, are tangibly connected with accident prevalence. The outstanding statistical proof of this fact lies in a study of the hours when accidents occur and of fatigue curves for night shifts. Vernon (1940) reports an increase in accident rate of two and one-half times during the latter part of the working day, approaching the period when light intensity is obviously at a minimum. Tiffin (1942) reports decreased efficiency in vision with increased age. He also reports employees with poor vision have more accidents both on and off the job than people with normal eyesight. There appears to be a definite relationship between accident rates of employees having a tendency to converge their eyes and those with a tendency to diverge their eyes, as determined on a distance phoria test. (See Figure 3.)

It is highly possible that such relationships as indicated in Tiffin's study may be the explanation of the increase in accident rate during periods of minimum light intensity, especially the increase in motor vehicle accidents at dusk and the higher accident rate experienced under artificial lighting conditions in industry.

WORKING CONDITIONS AND ACCIDENT RATE

Numerous studies completed under the auspices of the Industrial Health Research Board have firmly established relationships between accident frequency and working conditions. Farmer and Chambers (1926) distinguish between "accident proneness" and "accident liability," using "accident proneness" as a personal predisposition to accident and "accident liability" to include such items as hours of work, illumination, ventilation, temperature, and similar factors. Goldmark (1912) explains increased accident frequencies with progressive hours of work as the result of fatigue and its tendency to increase reaction time and cause muscular inaccuracies. Osborne and Vernon (1922) and Bedford and Vernon (1927) have presented studies indicating correlation between accident rates and atmospheric conditions of temperature, ventilation, and air movement. Stephenson (1926) attributes increased accident rate to inferior illumination. It is generally accepted that working conditions do bear a definite relationship to accident rate, and further discussions of the effect of these same factors will be found elsewhere in this volume.

GENERAL PREVENTION TECHNIQUES

Proper selection of personnel is the most effective method for prevention of ultimate accidents. Established techniques such as the administration of intelligence tests, personality tests, and dexterity tests provide readily available devices for determining whether the applicant meets the specific job requirements. Job analysis, preparation of adequate job descriptions, and restriction of employment to properly qualified personnel is the best method of conducting an accident prevention program. Training of the new employee is a vital safeguard, insuring familiarity with existing hazardous conditions, processes, and practices. Supervision with continuous observation and constructive criticism are most essential during the first month of employment, after which the new employee may be sufficiently oriented that there may be a relaxation of the supervision related to hazards. Maintenance of thorough accident records on all accidents, with an investigation of their cause, is essential to the prevention of recurrent accidents and adequate records and analyses of accidents by common factors of location, type, time of day, day of week, shift, job title, or other specific determiners will provide a source of factual information for analysis and determination of corrective action. Accident prevention must be treated exactly as any other industrial operation, for example, quantity control. The scientific method, applying Kipling's classical queries, is the fundamental basis of accident control—WHO—WHAT—WHEN—HOW—WHERE—WHY.

READING REFERENCES

Bridges, C. D., *Job Placement of the Physically Handicapped*. New York: McGraw-Hill Book Company, Inc., 1946, Chaps. 4, 5, 6; Appendixes B, C, G, H.

Brody, Leon, *Personal Factors in Safe Operation of Motor Vehicles*. New York: Center for Safety Education, Division of General Education, New York University, 1941.

Heinrich, H. W., *Industrial Accident Prevention*. New York: McGraw-Hill Book Company, Inc., 1931, Chaps. 2, 5, 8.

Maier, N. R. F., *Psychology in Industry*. Boston: Houghton Mifflin Company, 1946, Chaps. 8, 16.

Schaefer, V. G., *Safety Supervision*. New York: McGraw-Hill Book Company, Inc., 1941, Chaps. 4, 7, 9, 12, 14.

Tiffin, J., *Industrial Psychology* (2nd ed.). New York: Prentice-Hall, Inc., 1947, Chap. 11.

US Department of Labor, *Safety Subjects*. Washington, D.C.: Bulletin No. 67, Division of Labor Standards, 1944, Chap. 21.

Viteles, M. S., *Industrial Psychology*. New York: W. W. Norton & Company, Inc., 1932, Chaps. 16, 17, 18.

44. BUILDING ILLUMINATION

MILES A. TINKER, PH.D.

Professor of Psychology and Director, Lighting and Reading Laboratory, University of Minnesota; Consultant on Problems of Illumination, Typography, and Reading Efficiency.

Provision of adequate illumination is a paying proposition. From the point of view of individual hygiene, it assures greater comfort during visual work and promotes the conservation of eyesight. In the school, the office, the factory, and many other situations, satisfactory lighting yields more accurate performance and greater work output per unit of time.

NEED FOR BETTER LIGHTING

Most surveys of building illumination reveal marked deficiencies which preclude hygienic (i.e., healthful, comfortable, and efficient) vision during visual work. Koepke (1934) indicates that industrial lighting is often defective in both intensity and diffusion of the light obtained because it is prescribed without adequate planning. Frequently light intensity is only 2 to 5 footcandles when it should be much higher. In other instances illumination from poorly arranged fixtures is more intense than is desirable. Glare sources, which result in uneven distribution or diffusion of the light, are often prominent. All these conditions result in reduced visual efficiency, discomfort, and eyestrain which may eventually cause permanent damage to the eyes. Furthermore, poor lighting is considered to be a contributing factor in about 20 per cent of industrial accidents (Luckiesh and Moss, 1937). Many reports indicate a decrease in rate of accidents when the illumination has been improved.

A description of suitable illumination specifications that can be applied to all working environments is impossible in a brief discussion. Nevertheless, certain fundamental aspects of the problem can be briefly presented. In each situation, efficient lighting must be largely in terms of the specific visual task to be performed, especially as regards the brightness of the illumination. Visual efficiency is a psychological matter. For the most part only the physical facts of illumination are treated here; for their effects on individual efficiency see Section 10 of Chapter II.

THE NATURE OF LIGHTING

RELATION BETWEEN INTENSITY AND DISTRIBUTION OR DIFFUSION OF LIGHT

Adequate control of illumination distribution throughout the working environment is of prime importance. Any plan for lighting which aims merely to achieve a specified intensity is likely to be unsatisfactory. Hygienic vision is achieved only when brightness and distribution are adequately coordinated. Failure to maintain proper diffusion of light as intensity is increased produces lessened visual efficiency and eyestrain. This comes from the glare caused by poor arrangement and type of lighting fixtures, by surface reflection, and by an alternation of bright areas and shadows in the visual field.

Even with good distribution, illumination intensity can be excessively high. Publicity campaigns to improve sight by better light may have made the consumer willing to accept recommendations that are out of line with experimentally established requirements of hygienic vision. While somewhat higher intensities than are necessary may be employed with safety provided good distribution is maintained, the cost is greater and no additional gain in efficiency and comfort is achieved. Several pamphlets (American Recommended Practice of School Lighting, 1948; Recommended Practice of Office Lighting, 1942; American Recommended Practice of Industrial Lighting, 1942; Recommended Practice of Home Lighting, 1947) have been issued on lighting specifications. In these the discussions of lighting units, distribution of light, etc., are good. The suggestions on illumination intensities, however, should not be accepted at face value. The intensities specified in the recommended practices tend to be much higher than necessary for comfortable and efficient vision. A critical evaluation of the literature on illumination intensities in relation to visual efficiency is given in a paper by Tinker (1946).

TYPES OF LIGHTING EQUIPMENT

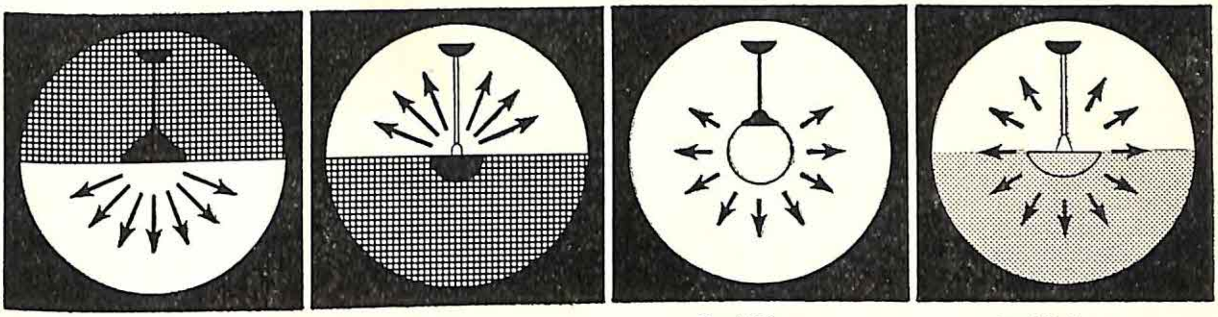
In general, there are four types of lighting units, classified with respect to the way the illumination derived from them is distributed. Direct, semi-direct, indirect, and semi-indirect units are illustrated in Figure 1. A fixture from which all the light is reflected principally downward and slightly outward produces direct illumination. Simple hanging lamps with an opaque reflector above, and bulbs with parchment shades are examples. Although effective in providing illumination on the working surface, it is difficult to avoid glare with this type of unit. In semidirect lighting a large proportion of the illumination comes directly to the working surface but part is reflected from the ceiling and upper walls. Most fixtures producing semidirect light are of the open-bottom type. Indirect lighting is characteristized by an arrangement in which all the illumination from the fixture goes to the ceiling and upper portions of the walls, from which the light is reflected diffusely about the room. A hanging fixture with an opaque reflector below the unit is an example. Semi-indirect lighting combines the features of the direct and indirect types. In this type some light passes through a white glass shade directly to the working surface. A large portion, however, goes to the ceiling and upper parts of the walls, from which it is reflected.

A uniform distribution of light can best be obtained with the indirect system. Glare is thereby avoided.

However, a compromise is usually made between efficiency, economy, and aesthetics. Too frequently this compromise results in unsatisfactory illumination which might be avoided by intelligent planning.

TYPES OF ILLUMINATION

There are, in practice, three kinds of illumination, i.e., general, local, and a combination of the two. With general illumination, the light from the units is spread to a greater or lesser degree about the whole room. With local illumination, the light is concentrated and is confined to a limited area of the working surface or workroom, as by a desk lamp with an opaque shade. Local lighting alone is likely to result in an unfortunate alternation of bright spots and shadows. Through a combination of moderately bright general illumination plus local lighting at the working surface one can achieve adequate intensity of light with satisfactory diffusion. The brightness of the general illumination should be greater than 10 per cent of the brightness of the central field where the visual work is to be done. In other words, the brightness ratio from local to general illumination should not exceed 10 to 1. A ratio of 5 to 1 is more desirable, and a ratio of 3 to 1 is ideal.



Direct Indirect Semi-Direct Semi-Indirect

Figure 1 (VI–44): *Types of Lighting Equipment*

Reproduced from *Lighting for Seeing in the Office,* by W. C. Brown and D. M. Warren, General Electric Company, Cleveland, Ohio, 1936.

COLOR OF LIGHT

Illumination from various sources has been used for visual work. Although sunlight contains wave lengths from all parts of the visible spectrum, the parts utilized by the eye in discriminating details lie largely in the greens and yellows, to which colors the eye is most sensitive. No light has been found to be more effective for seeing than daylight. The ordinary tungsten-filament lamp has a continuous spectrum. Its light is seen as yellowish, because of its spectral distribution. Daylight lamps are ordinarily equipped with blue glass bulbs; they emit a whiter light than the regular tungsten-filament lamp. In the mercury-arc lamp the electric current activates mercury vapor in the tubular bulbs. Although this lamp derives much of its light from the yellow and green sections of the spectrum, there is present some violet and blue. Its light is seen as bluish-green because of its discontinuous spectrum. The sodium-vapor lamp yields a truly monochromatic yellow light. Fluorescent lamps provide great freedom in the specification of spectral quality of light since various colors may be produced (white, blue, green, etc.). Furthermore, with this type of lamp, control of diffusion can be highly satisfactory and the radiant heat is relatively small. The tubular bulbs may be fitted into various types of ceiling fixtures and table lamps. The fluorescent lamps are basically mercury-vapor discharge tubes. The glass of these tubes is treated with fluorescent materials called phosphors. When activated by the ultraviolet light emitted by electrical excitation of the mercury vapor, these substances give off light in the visible spectrum. By appropriate choice of the phosphors, the fluorescent lamp can be made to emit light of almost any desired color. The operating cost of producing a given amount of light with the fluorescent lamps is relatively low.

Various experiments have demonstrated that visual efficiency is the same under tungsten-filament and under mercury-arc light (Luckiesh and Moss, 1937). The same is true for fluorescent light. With yellow sodium-vapor light, however, visual acuity is slightly better than with light from the tungsten-filament lamp. This difference, however, is so small that it has little practical significance. In general, the evidence (Tinker, 1939) indicates that one may employ whatever color of light he wishes for special effects and still maintain comfortable and efficient illumination, provided the factors of intensity and distribution are controlled.

LIGHTING ENVIRONMENTS

THE RELATION OF PAINT TO GOOD LIGHTING

Ceiling and wall surfaces are secondary sources of illumination since they reflect light received from lamps. Both brightness of color and finish are highly important. Bright colors with a high reflective factor (at least 75 per cent) are best for ceilings. White, ivory, and light cream are preferred for this reason. Somewhat darker colors of any desired hue may be used for the walls. A so-called flat or mat finish should be employed for both walls and ceilings because it disperses the reflected light without glare. Cleaning and repainting, when necessary, maintain the effective illumination.

LIGHTING IN THE HOME

Whether in the home, the factory, or any other situation where visual work is done, the principles of hygienic illumination discussed above should be adhered to. This implies proper coordination of intensity and distribution of light plus attention to the role of decoration. Certain details deserve further discussion. Intensity of light is measured in footcandles. One footcandle is the intensity of illumination from a standard candle at a distance of one foot. In reading, the light at the reading surface should be about 10 to 15 footcandles for books and about 15 to 25 footcandles for newspapers. For exacting visual tasks, such as sewing, or when one's vision is poor, there should be 30 to 150 footcandles according to the severity of the task (Tinker, 1946).

The home-owner or renter should not tolerate such glare-producing arrangements as wall brackets and low-hanging ceiling fixtures which provide annoying fatiguing light sources in peripheral vision. Also to be avoided is strictly local lighting, such as a desk lamp with an opaque shade, since this produces a bright spot of light surrounded by shadows. Table lamps and floor lamps should be tall enough so that the lamp shade is outside the field of vision, i.e., the bottom of shade should be 19 or 20 inches above top of table. All ceiling fixtures should be installed close to the ceiling. These fixtures should have relatively low surface brightness. While general plus local lighting may be obtained from ceiling fixtures supplemented by floor or table lamps, it may be achieved satisfactorily in a single fixture, the I.E.S. (Illuminating Engineering Society) certified lamp. This lamp has a diffusing bowl below the light bulb and within the shade. Thus some diffused light shines directly upon the working surface. Additional light goes from the open top of the fixture to the ceiling and upper walls from which it is reflected about the room. All lighting units should be so arranged that the reader or worker does not face the lamp. If the lamp is placed well to the side or behind the person, it will not be necessary to look at the light even when the eyes are raised from the work.

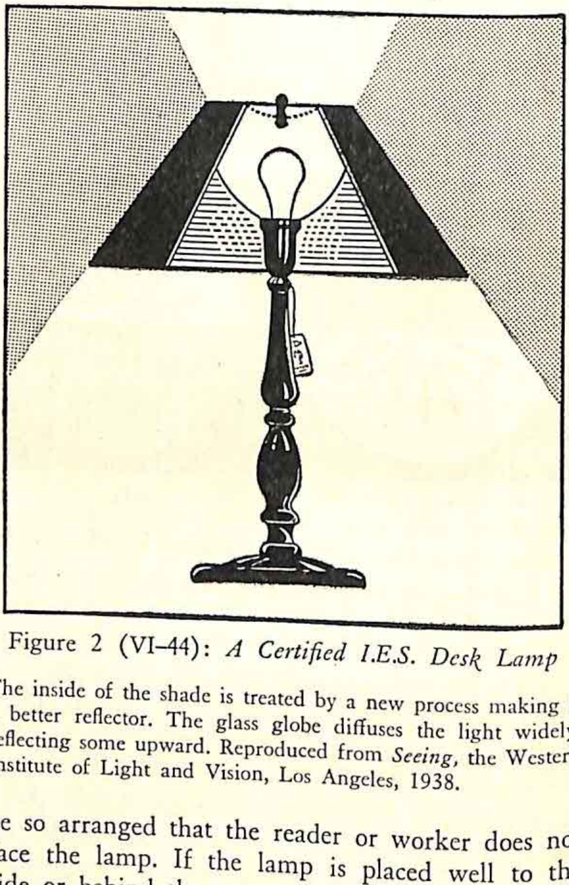


Figure 2 (VI–44): *A Certified I.E.S. Desk Lamp*

The inside of the shade is treated by a new process making it a better reflector. The glass globe diffuses the light widely, reflecting some upward. Reproduced from *Seeing,* the Western Institute of Light and Vision, Los Angeles, 1938.

A large variety of study lamps are available. In general the shade of a desk lamp should be outside the field of vision and the light emitted from the unit should be well diffused throughout the working area. The I.E.S. lamp illustrated in Figure 2 fulfills these requirements. Pin-up lamps that are similar in structure to the I.E.S. unit are fairly satisfactory. A fluorescent desk lamp with a fairly high standard is adequate if placed to one side of the worker, provided there is also general illumination of the room. Many desk lamps with opaque reflectors are inadequate when the lampshade is within the field of vision. Even though the light is diffused fairly well by certain of these units, there is still a sharp demarcation between the dark shade and the light below. Such a situation is improved by adding general illumination to the local lighting and by having a color of medium brightness (as neutral gray) on the outside of the shade.

LIGHTING IN LIBRARIES

In general, what has been said about home lighting applies to library lighting as well. There are, however, certain additional cautions and recommendations. Visual fatigue frequently arises because of the use of library tables with highly polished surfaces which reflect poorly distributed light into the reader's eyes. Tabletops should therefore have a dull surface even though beauty of finish is sacrificed. Also the tabletops should be light color to avoid brightness contrast with book pages. If permanent table fixtures with opaque shades are employed to supplement general illumination, as in many libraries, a more diffused light at the reading surface is obtained by either lumiline bulbs or fluorescent lamps than by the ordinary spherical bulbs. Both of these bulbs are long and rodlike, yielding well diffused illumination. This is especially true of fluorescent lamps. Intensity of illumination on the working surface should be adequate.

Directional lighting for getting illumination into drawers and onto upright books has a place in certain library situations, such as at card catalogue files and along bookshelves. Wiring circuits in reading rooms should be so arranged that the light units farthest from the windows can be turned on independently from those near the windows.

LIGHTING IN OFFICES

Clerical tasks in the office are more exacting visually than most of the activities carried on in home and school. The more severe tasks include drafting, reading shorthand notes, business machine operation, etc. The illumination should be 25 to 50 footcandles for many such office operations. In less exacting tasks, such as reading, writing, and mail sorting, 15 to 30 footcandles are adequate. The problem of controlling glare in large offices is by no means an easy one. Ceiling fixtures should, if possible, be maintained above the field of vision. The surface brightness of all ceiling fixtures should be low. General plus local illumination is often desirable. Fluorescent ceiling lamps facilitate adequate coordination of intensity and distribution of light. Indirect or semi-indirect lighting from silvered bowl lamps and other types of fixtures also may be used to good advantage in this respect. In all office lighting, maximum use should be made of ceilings and walls as reflecting surfaces. Fatiguing brightness ratios on office desks may be avoided by employing relatively light surfaces without polish.

LIGHTING IN THE SCHOOL

The International Commission of Illumination in 1931 suggested that the minimum light intensity in schoolrooms be set at 8 footcandles. More recent evidence indicates that about 20 footcandles is needed to provide hygienic lighting for all normal children, provided distribution is satisfactorily controlled. In sight-saving classes, however, there should be at least 45 footcandles. There are several sources of glare that should be avoided in schoolrooms. Desk surfaces should be light-colored and without polish to avoid brightness contrast and glare from surface reflection. Glossy paper should not be used as writing material. Also to avoid excessive visual contrast and glare, blackboards should be covered when not in use, and there should be no charts or blackboards between windows. Glare from daylight illumination may be controlled by venetian blinds or by translucent shades attached at the bottom or middle of the window. Photoelectric control devices may be employed for turning lights off and on to maintain the desired minimum of brightness. What has been already stated above concerning the coordination of diffusion of illumination with intensity and about the use of high reflecting, nonglossy finishes for ceiling and walls holds here also. Desirable usage in school lighting is well presented in *American Recommended Practice of School Lighting,* 1948, although the minimal intensities suggested are somewhat higher (Tinker, 1948).

LIGHTING IN FACTORIES

In most factories the situation is different from that in homes, offices, and schoolrooms, for little light is reflected from ceilings and walls to the working surface. Consequently lighting units are devised to direct most of the illumination downward. Fluorescent lamps are ideal for eliminating shadows and for providing adequate brightness of light in industry. When tungsten-filament lamps enclosed in globes are used for general illumination, additional local lighting is desirable for many operations. Data from earlier experiments suggest that 10 to 15 footcandles are sufficient for many factory tasks (Tinker, 1939). Where fine visual discrimination is involved, such as in certain types of inspection work, however, there should be 20 to 45 or 50 footcandles. The particular intensity needed varies according to the severity of the visual task (Tinker, 1946).

Maintaining hygienic illumination in the factory is profitable from the point of view of management. In addition to providing visual comfort for the worker, production increases with improvement of lighting (Tinker, 1939). Thus for six industrial operations, the original mean intensity of 2.4 footcandles was raised to 11.0 footcandles with an increase in production of 15.4 per cent. The accompanying increase in cost of lighting was relatively slight.

Two important sources of fatiguing glare are commonly found in factories. The first consists of bright light sources slightly above the worker's line of vision or to the side. In terms of reduced visual efficency, as much as 60 per cent of useful light may be wasted by these peripheral glare spots. Proper arrangement of lighting units will remedy this situation. A second possibility is reflected glare from polished surfaces, such as micrometers. This difficulty is eliminated by the use of well diffused light.

LIGHT IN CERTAIN OTHER SITUATIONS

In such places as hallways, auditoriums, and storerooms, where continued visual discrimination is not required, 5 to 10 footcandles of light is satisfactory. For surgical operations, where contrast is poor and need of good light is critical, several hundred footcandles are frequently used.

Inadequate lighting in buildings is inexcusable. Techniques are now available to achieve hygienic lighting in line with the fundamental facts derived from scientific research. Future trends in building illumination probaby will be along two lines. First, because the need for satisfactory light is receiving considerable publicity, the lighting requirements in a wide variety of situations will be more adequately met. More attention than formerly will be devoted to achieving a diffusion of illumination that will eliminate most of the preventable glare. Secondly, the present reaction against the use of the excessively high illumination intensities specified by engineers (Luckiesh, 1944) will be enhanced (Tinker, 1946). This will undoubtedly lead not only to less emphasis upon intensity as such, but to more stress upon the coordination of intensity with adequate diffusion of light.

READING REFERENCES

Illuminating Engineering Society, *American Recommended Practice of School Lighting*. New York: The Society, 1948.

Luckiesh, Matthew, *Light, Vision and Seeing*. New York: D. Van Nostrand Company, Inc., 1944.

Luckiesh, Matthew, and F. K. Moss, *The Science of Seeing*. New York: D. Van Nostrand Company, Inc., 1937.

Seymour, W. Douglas, *The Heating, Ventilation and Lighting of School Buildings*. London: Oxford University Press, 1939.

Tinker, M. A., "Illumination Standards." *American Journal of Public Health*, 1946, **36**, 963–973.

———, "Illumination Standards for Effective and Comfortable Vision." *Journal of Consulting Psychology*, 1939, **3**, 11–19.

45. INDUSTRIAL VENTILATION

L. P. HERRINGTON, PH.D.

Director of Research, John B. Pierce Laboratory of Hygiene, John B. Pierce Foundation and Department of Public Health, Yale Medical School.

Improved comfort, increased efficiency in work, and lowered accident and illness rates are the principal arguments in favor of control of indoor atmospheres (Vernon and Bedford, 1928). These points are well established. Although industrial and commercial ventilation includes such problems as odor control and removal of noxious gases or dusts, it is now recognized that the central problem of air conditioning and ventilation is the control of human heat loss. Bad air is not usually due to a depletion of oxygen or an increase in carbon dioxide, but rather to the unpleasant sensory effects and physiological sequelae of exposure to environments that place undue stress on the conservation or dissipation of body heat production. As F. S. Lee has aptly said, the problem is physical, not chemical; cutaneous, not respiratory.

PHYSICAL FACTORS

A dry-bulb temperature of 72° F, a radiant temperature of 68° F, a relative humidity of 60 per cent, and an average air movement of 15 feet per minute produces a definite cooling effect on the body. This cooling effect is a fixed physical value for a standard physical activity. In the experience of the average person, the cooling effect of an environment is associated with the reading of an ordinary mercury-in-glass thermometer. This dry-bulb temperature, however, is a measure of only one of the four important factors mentioned above. The meaning of dry-bulb temperature to the layman is based primarily upon experience with the large contrasts of indoor and outdoor temperatures in various seasons. For such ranges of cooling effect, it is the most important item. In the smaller range of comfortable or tolerable temperatures important in indoor conditioning, it is of far less importance. In the range from 68° to 75° F (dry-bulb) variations of air movement may be of greater importance, and above 75° F changes in relative humidity may be the leading factor in comfort. It is probable that the greater portion of public confusion on comfort conditions is directly related to this common misconception concerning the validity of dry-bulb temperature as a comfort index. For this reason, measurement of all of the four factors mentioned above is essential in understanding the problems of air-conditioning control. Uniform instrumentation is the remedy for this confusion and a satisfactory schedule has been prescribed by the American Society of Heating and Ventilating Engineers (Committee Report, 1939) for the measurement of the thermal factors from which the cooling effect should be computed. If the values for any of the four important factors are changed, the cooling effect is correspondingly altered. Obviously, various cooling effects are produced by different combinations of the thermal factors, even though dry-bulb temperature remains unaltered.

OPERATIVE TEMPERATURE

What we have been describing is called operative temperature (Gagge, 1940). It represents a method of arriving at a simplified expression for the combined cooling or heating effect of all atmospheric factors. Any environment in which air and wall temperatures are not equal, and in which air movement is not 15 feet per minute, has a physical cooling or heating effect on the body exactly equal to that which would be experienced in a standard environment. This standard environment is one in which air and wall (radiant) temperatures are equal and in which air movement is fixed at 15 feet per minute. Thus, for any possible combination of the above thermal factors, there is an equivalent standard condition of air and wall temperatures and air movement. The equal air-wall temperature of this equivalent thermal environment is the operative temperature. Example: Corrections for air movement only. Actual conditions: air and wall at 81.0° F, air movement 102 feet per minute. Resultant operative temperature, 71.6° F. (In a space with air and walls at 71.6° F and air movement at 15 feet per minute, the body heat loss would be equivalent to that experienced with an air-wall temperature of 81.9° F and 102 feet per minute air movement.) This scale is particularly useful where it is desired to bring environmental temperature into an exact relationship with the physiological temperature regulation of the body.

EFFECTIVE TEMPERATURE

Effective temperature (Annual Guide, American Society of Heating and Ventilating Engineers, 1948) expresses air temperature, air movement, and humidity in a single scale of psychosensory isothermals or just perceptible differences in warmth. It is valuable in locating comfortable conditions and in estimating the uncomfortable effects of increased humidities at temperatures above 80° F. The use of effective temperature to standardize different combinations of thermal conditions below 80° F is not encouraged since the scale does not correct for wall radiation, and probably overcorrects for humidity variations in this low range. In short, warm bodies do not cool at the same rate in environments which, on chart reference, correspond to a single effective temperature value. The best use of effective temperature is in forecasting the initial sensation of warmth or coolness on passing from one thermal condition to another. Where boiler surfaces, hot machines, or other sources of radiation are present, the absence of a correction for radiation is important. The effective temperature scale is based upon initial sensations of warmth rather than upon physical relations between human heat production and the cooling effect of an environment after prolonged occupancy. As a consequence of this fact, it cannot be used to calculate shifts of the comfort point which may be produced by changes of clothing or changes in body heat production arising from work or exercise. Such calculations are of great industrial importance, and advances made in these applications through the use of operative temperature will be discussed later. At temperatures above 80° F, Effective Temperature corrects for differences in humidity in a simpler manner than is possible with Operative Temperature. This correction may be applied, however, only when the conditions of work and clothing correspond to those used in the standardization of the Effective Temperature Scale. Operative Temperature is a more tedious calculation but is free of this restriction.

PHYSIOLOGICAL FACTORS

Thermal influences on comfort, efficiency, and health are certainly extensive (Lee, 1940), but their slow cumulative effect on the individual makes it difficult to isolate the effect from other hygienic factors.

Quite possibly the physiological cost of overheating is primarily circulatory. Pronounced overheating may elevate the heart rate by 40 to 50 extra beats per minute. Lesser degrees of overheating may be well tolerated subjectively, but there is no escape from the necessity of an added cardiac load. Cold environments may reduce circulation in the hands and feet to a value which is less than 5 per cent of the rate under comfortable conditions. In certain individuals this either leads to or exacerbates a tendency to serious vasomotor disorders of the extremities. Excessively dry environments of moderate warmth, particularly when experienced contra-seasonally, produce drying of the respiratory membranes, the outer skin, corneal surfaces, and hair, and in every group of reasonable size, certain individuals will suffer conspicuously from such effects.

The guiding principle in prescribing desirable environmental temperatures is to avoid abnormal stress on the heat regulatory mechanisms of the body, such as chilling, high evaporation, and increased circulation. This stress cannot be evaluated reliably from subjective reports. An accurate assessment of bodily stress involves a physiological examination in the thermal environment. This is important in industrial situations in which a compromise must be made between optimum comfort and desirable atmospheric conditions for industrial operations.

THERMAL GRADIENTS

The internal body temperature is approximately 99° F when at rest, normally clothed, and with average heat production. The internal heat must be moved to the periphery by circulatory activity, hence to do this the skin temperature must be below the internal body temperature. The same quantity of heat which is brought to the skin surface must be lost to the environment. Under average resting conditions, clothing surface temperature is about 8° F above the environmental temperature; the rectum is about 99° F, the average skin temperature about 92° F, the clothing and exposed skin about 83° F, and the air temperature about 75° F (Winslow, Herrington, and Gagge, 1940). In short, the gradient steps in the loss of heat from internal regions to skin to clothes to environment under comfortable conditions with sedentary heat production are 7° + 9° + 8° F.

THERMAL CONDUCTION CURVE

A view of the thermal relations between the body and the environment in terms of gradients is particularly convenient for industrial purposes. Figure 1 reproduces the thermal conductance curve for rectal temperature less skin temperature in centigrade. Values for tissue heat conductance in terms of large (kilogram) calories per square meter per hour are on the ordinate. Below are indicated the phases of bodily heat regulation associated with the various changes in the skin-internal gradient.

One should understand that in order for cooling to take place, heat must flow from a region of higher temperature to one of lower temperature. If the body is to maintain a constant internal temperature the heat which it produces must flow to a cooler skin and in turn to a cooler air. This is a condition of zero heat storage, indicated by $S = O$. Any increase in body heat production requires increased circulation of the warm blood through the skin to pass this extra heat over the same degree difference between the internal regions and the skin. Conversely, any increase in environmental temperature (tending to raise skin temperature) will require a like circulatory response even though the heat production is unchanged. In short, any change in heat production (rate of work), in environmental temperature, or in clothing produces circulatory or gradient changes in the body which are quantitatively interlocked. Hence, the importance in any thermal analysis of a study of the gradients in conjunction with knowledge of the level of heat production in the body. Small differences of 1° to 3° F between rectum and skin temperature mean that great circulatory activity is required to move body heat to the periphery; moderate differences of from 6° to 10° F indicate a minimal effort; and larger differences are associated with minimal circulation of blood in the skin and undesirable chilling. Optimal bodily regulation of heat occurs near the greatest inflection of the curve.

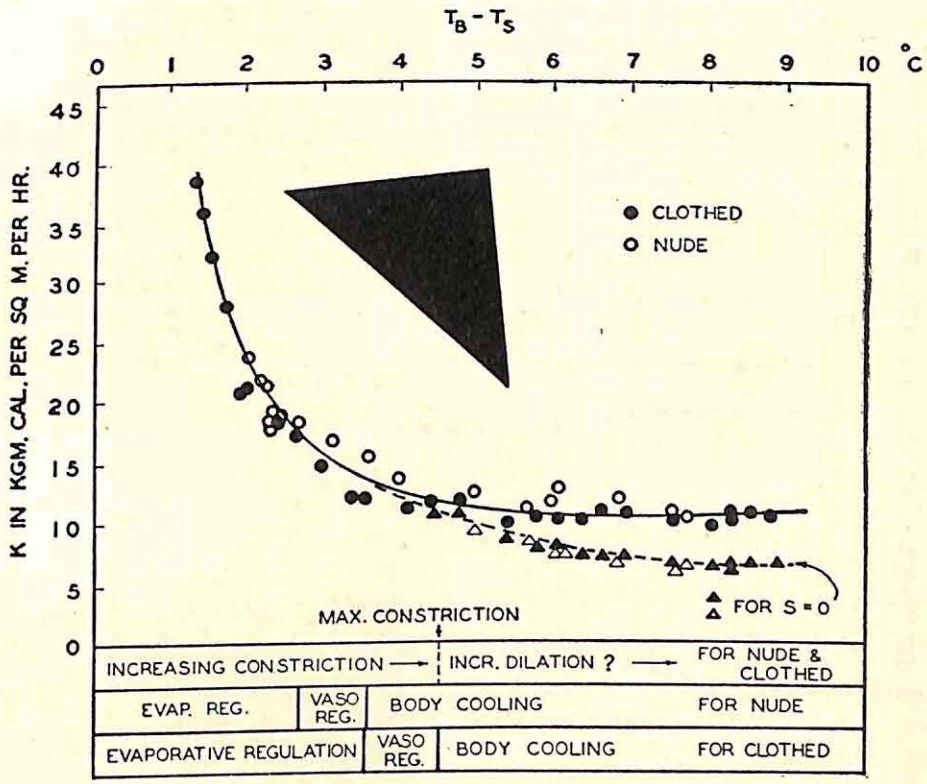


Figure 1 (VI–45): *Thermal Conductance Curve for Resting Young Males Showing Relationships between Temperature Gradient and Heat Conductance of the Body Tissues*

Solid curve shows observed values; dotted curve shows corrected values for a complete equilibrium of body and environment. T_B= rectal temperature; T_S= mean temperature of 15 unclothed skin areas.

The parameters of a bodily heat conductance curve, such as the one in Figure 1, are the relatively simple measurements of total heat production, rectal temperature, and average skin temperature. A similar curve may be plotted for any environment where man works and lives. The resulting data show the temperature range which places minimal stress on physiological functions with any type of clothing or level of work. They indicate in concrete terms the physiological cost of the temperature compromises which are often necessary between hygienic and industrially desirable conditions. The data of the heat conduction curve can be converted into a measure of blood flow. Both the blood output of the heart and the pulse elevations distinguished by Lee (1940) result in part from high temperature rather than from the increased metabolism associated with exhaustive work in hot environments.

HEAT BALANCE EQUATION

Increased heat production or alteration in environmental temperature produces a change in the gradients. The variations in gradients are included with other measures in a heat balance equation to estimate the effect of environmental temperatures on bodily activity (Gagge, Winslow, and Herrington, 1938). Figure 2 is the graphic presentation of the heat balance equation:

$$M + S = +R + C + E,$$

where the symbols represent metabolism, storage, radiation, convection, and evaporative heat exchange respectively (Gagge, 1940). Metabolism in this instance is the normal amount of heat produced by a body per unit of time. Storage is any temporary deviation from average bodily heat content as may occur when one goes into cooler or warmer places. Radiation, convection, and evaporation are external measures in calories of gain or loss in body heat by these physical processes.

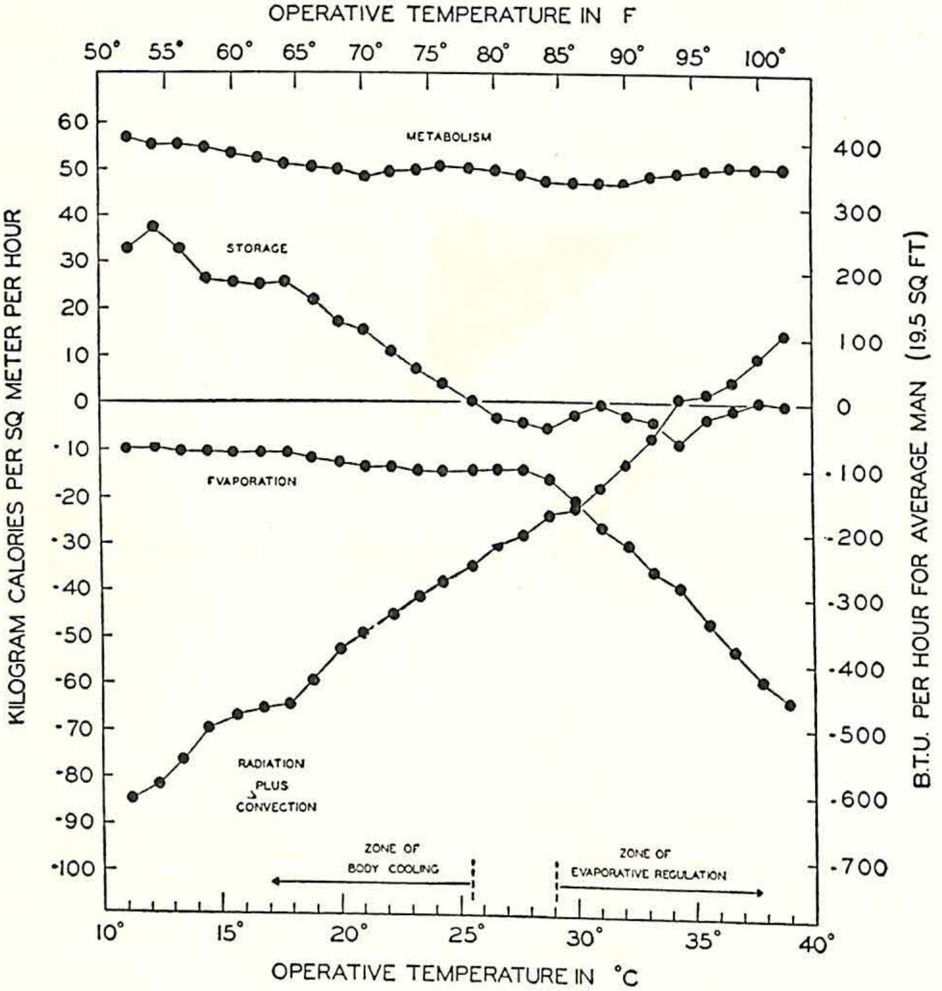


Figure 2 (VI–45): *Relation between Metabolism, Storage, Evaporation, Radiation plus Convection, and Operative Temperature for the Clothed Subject*

The ordinate of Figure 2 shows, for each process, the gain or loss in large (kilogram) calories of heat per square meter per hour. Physical operative temperatures, which are temperatures where air movement is 15 feet per minute, where wall surface is approximately at air temperature, and where relative humidity is from 50 per cent to 40 per cent, are shown on the abscissa. A vertical line connecting any four dots represents the heat balance of the body, according to the four measures of metabolism, storage, evaporation, and radiation plus convection, for these resting clothed men at any operative temperature indicated on the abscissa. These men produced 50 calories of heat per hour per square meter of body surface at 78° F, where there is little storage, of which 14 calories were lost from skin and lungs by evaporation, and 36 calories through body radiation and convection. Twenty-five calories were lost through these two sources at 85° F, where the evaporation curve shows increasing evaporation.

HUMAN TEMPERATURE REGULATION

There are three zones of human temperature adjustment, as illustrated in Figure 2: (1) The zone of body cooling by radiation and convection is from 50° to 75° F operative temperature. (2) The zone of comfortable balance by vasomotor control of skin circulation is the central zone which is approximately from 75° to 82° F. (3) The zone of uncomfortable balance by evaporative regulation is the region above 82° F. The average temperature of the body (not rectal) will be substantially different in the three zones. Storage represents the measure of this change.

ZONE OF RADIATION AND CONVECTION COOLING

Low environmental temperatures in the zone of body cooling by radiation and convection produce a constriction of the blood vessels of the skin. Heat loss is retarded and internal temperature is stabilized. Skin temperatures fall progressively with environmental temperature, and the person feels cool. The temperatures from 75° to 65° F are accompanied by a sense of stimulation or increased desire for activity. Lower temperatures are increasingly uncomfortable and the worker attempts to compensate for this by increased activity.

ZONE OF VASOMOTOR CONTROL

Heat balance is adequately controlled by blood flow to the skin in the region between the zone of excessive body-cooling and the zone of evaporative regulation. The worker feels indifferent, neither warm or cold. The tendency is toward relaxation rather than stimulation.

ZONE OF EVAPORATIVE REGULATION

Peripheral blood flow increases with increased temperatures and sweating increases in proportion to increased temperature. The loss of water in sweating is proportional to the amount by which environmental temperature exceeds normal skin temperature. The body is constantly stimulated by sensations of heat and there is an increased desire for relaxation. Work output is stabilized, as the limit of heat dissipation is reached, at a work level which is below the peak work load at lower temperatures. When a point is reached at which evaporation does not dissipate the body heat, skin and internal temperatures increase above the plateau which was established at the onset of evaporation.

HUMIDITY

Humidity is a factor of great importance above 80° F in influencing the heat regulative activity of the circulatory system. The ease of sweat evaporation increases with decreasing humidity at any given temperature. Table 1 gives the approximate temperature limits of evaporative regulation of body heat at various relative humidities. It shows the decline in air temperature necessary with increasing relative humidity if (1) mean skin temperature is stabilized between 95.9° F and 97.7° F, (2) the body surface is covered with perspiration (100 per cent wetted area), and (3) up to 1,000 grams of sweat can be physiologically produced per hour. The limits shown in Table 1 can be exceeded by well trained men capable of standing great discomfort. This is achieved, however, at the cost of severe circulatory effort and elevated body temperatures. It should be stressed that these combinations of air temperature and relative humidity are not approved conditions for unpleasant but necessary exposures. On the contrary, they represent severe conditions which may be expected to produce serious distress in a considerable part of the population. Any condition exceeding the limits noted will require a selected personnel and close supervision.

Table 1: Temperature Limits of Evaporative Regulation

	Environmental Temperature in °F			
Relative Humidity in Per Cent	Moderate Work	Resting Nude	Resting Nude Air Movement*	Resting Clothed
0	120.4	117.7	141.0	126.9
20	108.0	109.2	121.3	111.9
40	99.7	102.7	111.0	103.3
60	93.9	98.1	103.6	97.0
80	89.6	94.1	98.4	92.5
100	85.4	90.7	93.9	87.8

* 100 feet per minute; all other figures apply to low velocities (15 feet per minute), commonly referred to as still air.

Air movement raises the temperature limits of evaporation given in Table 1. Moderate work reduces the temperature limit, particularly in the higher humidity range. Clothing extends the temperature limit at zero humidity. The rather surprising clothing effect is well known to the Arabs and other desert peoples. At air temperatures above skin temperature, the clothing impedes the reception of heat from the environment and favors evaporation through a *wick-action* of the clothing. This results from the fact that all sweat secreted may be dispersed in the clothing and rapidly evaporated, thus reducing the dehydration stress which results when the semiclothed body loses the cooling effect of a certain portion of its sweat through dripping.

CONDITIONS OF EFFICIENCY

Comfort is of real importance in determining optimum atmospheric conditions in sedentary work, as atmospheric conditions for greatest efficiency and maximum thermal comfort probably are identical. Even if comfort and efficiency were not closely related, no worker willingly tolerates uncomfortable conditions for the sake of efficiency. Hence the determination of comfortable working conditions is the essential task of the ventilation psychologist. A wide variety of mental tasks have been performed with equal efficiency under cold, hot, dry, and wet conditions, as well as those just comfortable (New York State Commission on Ventilation, 1923). But these experiments were performed in a laboratory under highly motivated conditions. Efficiency falls off in all practical situations under cold conditions, probably because of distraction effects; and lowered motivation under hot conditions produces a definite decrement in performance (Huntington, 1925).

OCCUPATIONAL INFLUENCE

Comfort is determined by the task as well as the temperature. Workers normally clothed and required to sit quietly for several hours consistently selected 77° to 80° F as an area of maximum comfort (Gagge, Winslow, and Herrington, 1938). Occupied subjects similarly clothed but with no important increase in heat production prefer 70° to 72° F. For a clothed subject not physically active, 75° to 80° F is the range within which the heat regulatory reflexes change from skin vasoconstriction to vasodilation. The unoccupied subject is comfortable and less bored, apparently, if his temperature regulation is in a phase of dilation. In short, if you have nothing to do, stay on the warm side and you will enjoy it.

MENTAL WORK

Subjects required to do mental work, reading attentively, or writing, select 70° to 72° F as the ideal temperature. At this temperature they are slightly vasoconstricted, which is a condition favoring an attentive rather than a relaxed mental attitude. This fact is of some importance in the school situation. Women teachers, by reason of sex and age, are likely to experience maximum comfort at a point several degrees above that desirable for children. Much the same situation occurs in libraries and other public institutions in which temperatures are regulated largely by older individuals with slightly reduced heat production and less efficient circulation in the extremities. The practical rule for mental or sedentary work is to keep environmental temperature slightly in the zone of body-cooling. Temperature conditions for physical activity should make only minimal demands on physiological heat regulation. Acceptable standards exist for sedentary and light work (*Annual Guide,* American Society of Heating and Ventilating Engineers, 1948). It is proper to emphasize the flexibility of "comfort zones" as a function of activity, types of clothing, sex, and age.

COMPLAINT SITUATIONS

Any prescription of desirable atmospheric conditions in industry usually involves groups of workers, and in every group there are conspicuous exceptions to the average for whom the prescription is made. The psychologist not infrequently contacts problems of industrial efficiency and comfort resulting from a too rigid application of ventilation prescriptions made for groups of workers.

SPECIFIC THERMAL FACTORS

It is not uncommon in industry to use dry-bulb temperature alone as an indicator of the total atmospheric situation. Air temperature, however, is only one thermal comfort factor. A complete analysis must be made of all specific thermal factors, such as humidity, hot or cold radiation effects, and air movement. It is necessary to do this in order to effect changes which may remove complaints resulting from a peculiar combination of thermal factors which, because of dress or room location, affect certain individuals more than others. The measures of Table 2 are recommended for this purpose. They will provide an air temperature, air movement, and radiation map in sufficient details to permit the analysis of individual complaints. Air temperature and Globe thermometer readings (radiation effects) should be made in the horizontal plane, and positions should be selected such that a set of readings is available for approximately every 75 square feet of floor surface. Air conditions are usually controlled from a single point in the room, and deviation from a central thermostat reading throughout the room space may be very important in the comfort of the workers. Provisions are made for this in the measurement. There are likely to be substantial differences between center and periphery, locations near inside and outside walls, and between foot and head level. There is a frequent relation between individual position in a room, particularly with regard to head and floor level temperature differential and individual complaints. Objectionable drafts are noted when the velocity of an air current is more than 40 feet per minute, and the temperature of the air current is more than 2° F below the customary winter room temperature. Drafts are often reported when the back or head is exposed to direct low temperature radiation from a cold surface, such as a single glazed window or skylight.

Table 2: Measures Necessary for Thermal Analysis of Space Occupied by A Group

Level of Air Temperature*	Level of Globe Thermometer Readings (30-inch level)**	Level of Air Movement*	Relative Humidity
6-inch 30-inch 5-foot Ceiling	Room center Corners, and between occupants and local hot or cold surfaces	6-inch 5-foot	In zone of occupancy

* For each 75 square feet of floor space
** A blackened copper sphere 6 inches in diameter with ordinary thermometer inserted with bulb at center of sphere

Uncomfortable atmospheric conditions affecting the individual worker result from (1) special effects of radiation, air movement, or humidity, (2) errors in thermal control due to assuming that measurements at a control point are representative of an entire space, and (3) errors due to neglect of differential temperature between head and feet of workers. The survey recommendations of Table 2 are designed to give the information necessary in checking these complaint factors.

BODY HEAT

Engineers frequently misinterpret the role of human heat production in ventilation situations and think of resting heat production values as standard. Average figures for the heat production associated with various activities and occupations are listed in Wiggers (1934). If the heat production is doubled or tripled by work, the new comfort condition may be roughly estimated by subtracting from the normal sedentary comfort level of 70° to 72° F, a degree gradient sufficient to dissipate the added heat. With heat production doubled, one may lower the temperature from 4° to 7° F; if tripled, from 8° to 14° F. The higher degree value is applicable where the work is largely static, involving only moderate air movement due to self-motion.

AIR MOVEMENT

Increased air movement is usually thought of as related to increased cooling and increased clothing associated with increased warmth. This is not always true and application of these opinions may result in the opposite effect in certain situations. The human body stabilizes its skin temperature at about 95° F when sweating begins. As air temperatures rise above this, increased air movement will not remove heat from the cooler skin or clothing by convection, and cooling will be increased only insofar as increased air velocity increases evaporation at a rate greater than the heat *gain* by convection from the hot air. Similarly, the upper limit of human tolerance to air heat is actually increased by normal clothing at low humidities, a result attributable to the greater cooling by sweating when the clothes are available as large surface wicks. This effect is seen in Table 1, which shows limits of heat tolerance for sedentary activity, with different humidities for both the clothed and the nude body.

INDIVIDUAL AND SEX DIFFERENCES

It seems quite impossible to set up ventilation conditions which will be entirely satisfactory for all members of a group, even though they are of comparable sex and engaged in the same activity. This is conspicuous in sedentary or very light activity. It is the result of differences ranging from ±10 to ±15 per cent in individual heat production during the same activity, individual peculiarities in clothing, different individual susceptibility to temperature sensations, and different levels of activity carried on in the same enclosure. By far the greatest difficulty in establishing satisfactory ventilation for the individual members of a mixed group arises from the different clothing habits of men and women. The clothing worn indoors by the male weighs from 6 to 7 pounds and is decidedly protective. Clothing worn by women under similar circumstances weighs 1 to 3 pounds and is of such materials and so disposed as to be much less effective in preventing heat loss. Since women normally have lower heat production than men, it is not surprising to find that optimum comfort is practically impossible for members of both sexes in the same room. There is no solution to this problem in industry if clothing habits remain as they are at present.

ABNORMAL RESPONSES

Physiologically normal individuals are not noticeably affected by either high or low humidities over a wide range of moderate temperatures, 50° to 75° F. However, certain individuals, frequently of constitutionally asthenic build, are predisposed by lower winter temperatures to excessive peripheral vasoconstriction. Skin secretions often are deficient and there is a drying effect on the skin surfaces. Eyes, hair, and nasal membranes frequently are similarly affected. These individuals seem to be especially susceptible to low humidities, possibly as a result of deficient skin hydration along with chronic vasoconstriction during the cool season. They are more commonly females and it may be that their lighter clothing increases ventilation of the skin and a more decided drying effect. Increase of relative humidity to 50-60 per cent, which does not cause normal workers to feel uncomfortable, or increased clothing for the individuals affected is the best solution of this problem. The afflicted frequently ascribe their discomfort to furnace gas, vitiated air, or obscure industrial irritants, which are usually not present. But such complaint, if not dealt with properly, can cause a considerable morale problem in a working force.

THE INDUSTRIAL PSYCHOLOGIST AND VENTILATION PROBLEMS

Technical facts and principles contributed by the psychologist, the physicist, the physiologist, and the engineer are necessary to a satisfactory solution of problems of ventilation and thermal hygiene. The psychologist is well fitted to assemble this knowledge for application in the industrial situation since the ultimate criterion is the comfort and efficiency of the industrial worker. One of his most important problems is the organization of his findings in terms which are readily grasped by executive personnel unfamiliar with the details of the problem. Great progress was made during the war in developing a terminology and testing routine which could accomplish this end (See Gagge, Burton, and Bazett, 1941). A unit of clothing insulation, the "Clo," and of heat production, the "met," were defined in terms of various familiar clothing assemblies and types of working activity. These practical units were applied to the testing data obtained from artificially heated mannikins that were constructed of copper or plastic. Their heat production may be varied to correspond to any level of human work. They may be clothed in various working assemblies and exposed to the atmospheric conditions obtaining in the work environment. The skin temperatures of such models afford a ready index of the effect of activity, clothing, or environmental variation. When used in conjunction with the physiological data on heat conductance (Figure 1) and psychological reports on comfort, they afford a graphic illustration of the thermal problem presented by any given work situation. By proper reference to the existing literature, these results may be interpreted by the industrial psychologist in terms which relate discomfort to a tangible physiological cost. In addition, the conditions which produce the stress, and those required to remedy it, can be demonstrated by the use of heated models in a manner which seldom fails to produce understanding and engage the interest of responsible officials. Although most of these techniques are not conventional psychological procedures, the end result, worker comfort and efficiency, is the decided province of the industrial psychologist.

The techniques discussed here represent a useful addition to the more familiar procedures developed by the psychologist in connection with other problems of the industrial environment.

READING REFERENCES

American Society of Heating and Ventilating Engineers, "Committee Report on Physiological Factors." *American Society of Heating and Ventilating Engineers Transactions,* 1939, **45**, 111–112.

———, *Heating, Ventilating, Air Conditioning Guide, 1948.* New York: The Society, 1948, Chap. 12, "Physiological Principles."

Gagge, A. P., "Standard Operative Temperature, a Generalized Temperature Scale Applicable to Direct and Partitional Calorimetry." *American Journal of Physiology,* 1940, **131**, 93–103.

———, et al., "A Practical System of Units for the Description of the Heat Exchange of Man with His Environment." *Science,* 1941, **94**, 428–430.

———, C. E. A. Winslow, and L. P. Herrington, "The Influence of Clothing on the Physiological Reactions of the Human Body to Varying Environmental Temperatures." *American Journal of Physiology,* 1938, **124**, 30–50.

Lee, D. H. K., *A Basis for the Study of Man's Reaction to Tropical Climates.* University of Queensland, Department of Physiology Publication No. 5, 1940.

Vernon, H. M., and T. Bedford, *The Relation of Atmospheric Conditions to the Working Capacity and Acci-*

dent Rate of Coal Miners. London: Industrial Fatigue Research Board Report, No. 39, 1927.

———, T. Bedford, and C. G. Warner, *A Study of Absenteeism in a Group of Ten Collieries.* London: Industrial Fatigue Research Board Report, No. 51, 1928.

Wiggers, C. J., *Physiology in Health and Disease.* Philadelphia: Lea and Febiger, 1934, Chap. 57, "The Material and Energy Balances of the Body, Nitrogen Equilibrium, Basal Metabolism."

Winslow, C. E. A., L. P. Herrington, and A. P. Gagge, "The Reactions of the Clothed Human Body to Variations in Atmospheric Humidity." *American Journal of Physiology*, 1938, **124**, 692–703.

46. INDUSTRIAL MORALE

FRANCIS F. BRADSHAW, PH.D.

President, Richardson, Bellows, Henry and Company. Formerly: Expert Consultant, Personnel Research Section, AGO, Department of the Army; Dean of Students, Director of the Bureau of Vocational Information, Director of the University Testing Service, and Dean of the College for War Training, University of North Carolina.

HERBERT E. KRUGMAN, BSS.

Staff Member, Richardson, Bellows, Henry and Company. Formerly: Personnel Consultant, Psychological Branch, Office of the Air Surgeon, Army Air Forces.

"Morale" is a word variously and somewhat loosely used to describe the general quality of the relationships among members of a team or group that radically influence the group's effectiveness as a working unit.

In investigations of civilian and military morale during the recent war ("What Is Morale?" by the Morale Services Division, 1943), the *level* of morale was described as being dependent on or composed of such elements as the following:

Faith in the cause and its future.
Pride and confidence in one's own group.
Belief in one's mission.
Confidence in training and equipment of one's unit.
Realistic appraisal of the job ahead.
Satisfaction with one's individual job assignment.
Belief in the leader's concern for one's individual welfare.

Such an analysis as the preceding is largely verbal. It is not clear that social psychology has yet been able to isolate independent variables of status and motivation that add up to what we mean by the term "morale." Yet it seems clearly profitable to retain the term until we find a better one. Certainly any team, committee, social institution, or business organization is characterized by something in addition to the specific job proficiencies of its various members. The readiness with which the individuals undergo hardship for the welfare of the group, their eagerness to render mutual aid in accomplishing a specific job assignment, the enthusiasm and determination that inspire and support their daily efforts to solve problems and maintain schedules—these may be even more important in determining the output of a work group than the degree of talent and training represented in the group.

Earlier students have tended to agree that industrial groups are so constituted that the degree of cohesiveness and team work seems to depend, in part at least, on the extent to which the various individuals have:

A feeling of common ownership.
A sense of economic security.
A belief in the fairness of competition with other members of the group.
A knowledge of progress toward individual and group goals.
A feeling of assurance that they are accepted in the group.
A belief in the ability and justice of top management.

MEASUREMENT

Further progress in the scientific study of industrial morale requires increasing precision in definition and the development of instruments for measuring the level of morale and the changes produced by known modifications in the factors that determine it. To date, three main methods of measuring morale have been used in industry, namely:

Opinion polls.
Participation measures.
Multidimensional attitude studies.

Opinion polls are usually based on paper and pencil questionnaires administered to groups of employees in order to find out their opinion of management policy, working conditions, supervisory practices, and other specific subjects thought by management to be of importance at the time. Such polls provide much useful information on a relatively superficial level. While they help to correct many management conceptions about "what the workers think," they actually reveal, not exactly what the workers think, but what they say they think when asked in this limited fashion. What the workers really think or feel or will do may not be entirely indicated by the results of an opinion poll.

Since management's real concern is to learn to what extent morale can be improved, and how, it is usually important to obtain a more thoroughgoing knowledge of the social forces operating in the group than can be ascertained from an opinion poll. In order to do this, renewed efforts have been made to measure morale by the acid test of behavior. The question then becomes, "How do the individuals participate in the life and work of the group?" Absenteeism, turnover, efficiency measures, attendance on group occasions, activity in making suggestions, spoilage rate, accident rates (both lost time and non-lost time), tardiness, housekeeping neatness, applications for employment, volunteering for new training, and other such objective data are analyzed and collated as a chart of the rise and fall in morale.

During recent years, especially since the Western Electric experiment at Hawthorne (Roethlisberger and Dickson, 1943), emphasis has been placed on multidimensional interviewing and on finding out not so much what worker opinion is, but why and how it came to be so. Such a method differs from the opinion poll in that it depends on a more thorough treatment of small but statistically sound samples of employees. Anywhere from 1 to 10 per cent may be used, depending primarily on the total number of employees involved. However, since the sample must be representative of all levels, locations, and occupations within the plant, the diversity or homogeneity of these factors will also have to be considered in determining sample size. One current version of the multidimensional attitude study involves three waves of individual interviews with half or more of the total sample of employees being interviewed in the second or middle wave and the balance distributed equally between the first and third waves. The first wave is primarily open-ended and nondirective. It is devoted to the identification of the most troublesome problems. The second wave is focused more sharply and involves intensive exploration of problems and searching for cause and effect relationships. The third wave is concerned with checking and clarifying the results of the second. Each wave of interviews, therefore, is hooked up with the one following in a careful, evolutionary sequence. Shortly thereafter larger groups may be queried by a carefully constructed questionnaire so that the reliability of the interview sample and the generality of the conclusions derived from it may be more rigorously checked. At the same time, general conditions of workmanship, efficiency, absenteeism, and other objective data also are analyzed. A detailed description of such procedures may be found in a report by Campbell (1948).

The survey team will usually represent a high degree of professional competence in psychology—possibly clinical and social, as well as industrial. The synthesis and interpretation of results will normally be done by a team of still greater competence, in conjunction with the interviewers, and will require the additional services of the engineering and accounting departments. Finally, a report will be made to management (usually also to the employees) attempting to specify from these data where and why unnecessarily low morale has been found. Specifically, a prediction is made as to *how* such morale can be raised. In this type of measurement there is relatively little interest in comparing one company with another. The question is mainly, "What specifically can be done to improve this particular patient's health?"

THE PROBLEMS OF INDUSTRIAL MORALE

Commercial and industrial organizations are made up of many varied groups—owners; employed managers; professional, technical, and research employees; middle-management; first-line supervisors; and unionized and nonunionized rank and file. In addition to these divisions, roughly by level, there are, in many large organizations, divisions by function—sales, production, accounting, administration, research, and the like. The morale of any subgroup is dependent, naturally, on the extent to which its own peculiar objectives are central among the objectives of the entire organization.

As the goals of a total organization become less strong and less specific, the goals of subgroups and individuals increase in importance as morale determiners. In extreme cases there will be groups that cannot maintain a high level of morale except by satisfying all the basic goals of individuals and subgroups. At another extreme, organizations will have such high morale that, for a time, individuals and subgroups actually forget their own welfare as individuals, that is, their partisan interests. At present, however, most commercial and industrial organizations have a morale structure that makes the goals of individuals and subgroups extremely important. This is because the owners and investors are frequently absentee owners. Management itself may have only infrequent and vague contacts with the owners and, in large organizations, equally infrequent and vague contacts with the rank and file.

The first step that the managers of this type of industrial organization must take in order to handle their morale problems more effectively is to recognize the multiple-goal nature of the employed group that makes up its organization. Managers have been slow to do this because they have been closer to the owner-investor group in their concern for, and personal share in, the attainment of economic goals. In evaluating the behavior of lower-ranking employees, they have, very naturally, overemphasized the singular importance of these goals. Reaction to this overemphasis has been pioneered most ably, in the social sciences, by Elton Mayo's vigorous campaign against what he called the false theory of "economic man" (1933, 1945), which culminated in the extensive industrial researches referred to as "The Hawthorne Study" (Roethlisberger and Dickson, 1943).

Industry has undoubtedly made significant progress toward awareness of the noneconomic factors involved in morale, but the advance from the more to the less tangible of these goals—for instance, from an awareness of the importance of physical working conditions and fatigue to, at the other end of the scale, an awareness of such factors as self-respect and the feeling of belonging—has been slow. Most industrial organizations have apparently not reached even the half-way mark in awareness. Some, of course, are fully cognisant. Labor organizations, too, are recognizing more and more the less tangible goals of their members; the study of Golden and Ruttenberg (1942) may be one indication of this trend.

We may turn from the strength of group purposes to their complexity. The complexity of the over-all group purpose will have direct implications for the number of cooperating subgroups that must be formed to carry out successful realization of the group's over-all purpose. In our modern industrial society, with its highly specialized division of labor, this type of situation is very common. In most industrial organizations, the over-all goal of the group is not very strong for the great majority of these employees. This factor, plus the separation of subgroups in terms of type of work done, tends to make individual identification with the subgroup (shop, office, department) more realistic, more immediate and satisfying, than identification with the larger group. We then have a whole series of groups with a coordinating (management) group above them.

When the situation just described is characteristic of an organization, several difficulties arise. Since the subgroups have now taken on many of the characteristics of individual groups, there may be great variation in both the level of morale existing in each of the subgroups and in the local determinants for each of those levels. This means that technical coordination between the subgroups may fall down because of low morale in any one of them, and the achievement of the over-all goal thereby delayed. Thus a much greater degree of responsibility rests with local leadership in each of the subgroups, both in terms of supervision of personnel, and in terms of facilitating communication up and down between management and the rank and file.

The specialized coordinating functions of management have, in the larger organizations, come to require a degree of skill acquired only after long experience and gradual advancement through the ranks, or concentrated academic training in administration skills. The management group is therefore largely composed of people recruited from outside the plant or of people who usually have not, for many years, been members of the rank and file. Times and positions change, personal contact fades away, and, amid the whirl of technical duties, the top and bottom of the organization develop pictures of each other distorted by differences in time, personal experience, and individual satisfactions on the job. Unless this gap is bridged by improved communication, the coordinating group loses the opportunity to make progress in three different ways: (1) it is less able to draw the subgroups together into the larger over-all group and thereby make the over-all group purpose stronger for the rank and file; (2) it is less able to handle morale problems of one common pattern, but must expend far more time and energy attending to varying patterns in different groups; and (3) it is less able to obtain accurate facts on which to formulate the solutions of its morale problems.

In view of the multigoal nature of our typical modern industrial group, the next section will be devoted to an outline of the many individual goals that operate instead of the single all-embracing industrial goal previously pointed out as unreal. The last section will consider the prominent obstacles to these goals and how to surmount them.

THE INDIVIDUAL FACTORS

The most recent published attempt at classifying basic human needs in industry was made by Bakke (1947). On the basis of extensive interviewing with rank and file members of industrial organizations, he states, "Men want to experience security, progress, and justice with respect to the attainment of the following goals. . . ." These are listed below:

1. Respect of fellows.
   The desire to play a socially respectable role.
2. Creature sufficiency.
   The desire to have that degree of creature sufficiency (food, clothes, shelter, health, and the like, and the means to provide them) enjoyed by the most favored of one's customary associates.
3. Increasing control over one's own affairs.
   The desire to have one's decisions and actions effective in shaping the course of one's own life and to reduce the control exercised by others.
4. Understanding.
   The desire to comprehend the forces and factors that operate in one's world, to "know the score."
5. Capacity performance.
   The desire to utilize in living the full range of capacities possessed, both actual and potential.
6. Integrity or wholeness.
   The desire to experience consistency within one's self, among the parts of one's world, and a significant relationship to that world.

Although these goals are not scientific entities, they are descriptively convenient for practical use. They certainly offer a challenge to any management group sincerely concerned with the achievement of optimal employee morale. Lest the challenge seem too great or the goals too vague, it might be well to include another descriptive list of a somewhat different nature.

In connection with their activities in the field of industrial human relations programs, the authors have worked with discussion groups composed of two representatives each from six or seven management levels. In one manufacturing plant, after meeting once a week for twenty weeks, such a group spent three successive meetings developing its own list of basic human needs and translating each of these broad needs into the specific characteristics and practices of its own local industrial situation. The list is similar to those developed by academicians. In this case, however, the detailed sublistings offer a vivid illustration of what the major factors mean to practical businessmen. The list developed by this discussion group is presented below, without modification.

I. FOOD AND PHYSICAL WELFARE
   1. Working conditions—dust, light, arrangement of tools, noise, fumes, temperature, moisture
   2. Shelter
   3. Transportation
   4. Living wage—wage sufficient for minimum physical requirements
   5. Food during working hours
   6. Facilities—lockers, etc.
   7. Clothing
   8. Rest periods
   9. Freedom of motion—opportunity to change physical position and relieve muscular fatigue
   10. Housekeeping—cleanliness of surroundings

II. REPRODUCTION (HOME MAKING)
   1. Housing—opportunity for home making
   2. Social recreation
   3. Living wages—family requirements
   4. Working hours and schedule (hours at home)
   5. Transportation (hours at home)
   6. Community opportunities for children
      a. Education
      b. Social groups
      c. Church activities
   7. Confidence in the future

III. DESIRE FOR ACHIEVEMENT OVER DIFFICULTIES
   1. Having a known goal
   2. Knowing what is expected of an individual
   3. Having a measure for individual accomplishment
   4. Having available proper training to achieve what is expected of the individual

5. Having a clear understanding of the limits of an individual's job
6. Having a proper reward as recognition for desired achievement
7. Guidance into activity appropriate with individual ability
8. Making provision for steady progress

IV. DESIRE FOR ACTIVITY, VARIETY, AND NOVELTY

A. *On the Job*
   1. Determination of proper limits of job
   2. Opportunity for movement
   3. Relief from monotony
   4. Rest periods
   5. Recognition of individual differences
   6. Change from job to job—with limits
   7. Change of job within group
   8. Opportunity for initiative in thought

B. *Off the Job* (Company and non-company sponsored)
   1. Having proper recreational facilities
   2. Social activities

V. RELEASE FROM EMOTIONAL TENSION

1. Opportunity for physical activity
2. Having proper grievance outlets
3. Having clearly understood policies particularly regarding status
4. Counseling—listening
5. Social contacts on and off job
6. Religious activities
7. Security
8. Proper organization and selection of working group and supervision
9. Enough properly trained supervisors, sympathetic, easily available
10. Good home atmosphere
11. Confidence in fairness of wages, promotion, etc., task assignment
12. Belief in justice of management
13. Smoking on rest periods
14. Miscellaneous reducers like gum and candy

VI. SECURITY OF STATUS

1. Security of employment and income
2. Pension systems
3. Seniority
4. All types of insurance
5. Definite and understood promotion policy
6. Company stability
7. Company reputation
8. Good company discipline
9. Impartial treatment of employees
10. Freedom from coercion from unions
11. Stability of purchasing power
12. Confidence in management
13. Opportunity to maintain personal habits
14. Confidence in social order

VII. WORTHY GROUP MEMBERSHIP

1. Recognition of a good job (performance rating)
2. Belonging to a good company
3. Chance to join safety committee, fire brigades, first aid
4. Acceptance by group
5. Union membership
6. Group pride or morale
7. Inter-group competition
8. Congenial grouping
9. Company-wide occasions and symbols

VIII. SENSE OF PERSONAL WORTH

1. Democratic disciplinary system
2. Recognition of individual effort
3. Execution of suggestions made
4. Being consulted
5. Being treated as a person
6. Having his opinion treated with respect
7. Being made to feel important to supervisor
8. Being made to feel important to members of his group
9. Being called by name
10. Sympathy toward individual problems of others
11. Recognition of religious feeling
12. Having absolute value as a unit of life
13. Having a feeling of worth-while accomplishment

IX. A SENSE OF PARTICIPATION IS INDUCED BY THESE FACTORS

1. Knowing what is going on
2. Knowing why it is going on
3. Knowing the use of end products
4. Helping determine the conditions under which he works
5. Knowing the relation of his job to the finished product
6. Knowing the relations with fellow workers
7. Knowing the relation of his group to other groups
8. Helping in the administration of policies
9. Knowing relation of his job to supervisor's job
10. Having knowledge of future plans
11. Knowing how outside conditions affect his job
12. Sense of citizenship in the community

13. Knowing the relation of his job to others in his group
14. Opportunity to participate in group recreation

SOME MAJOR OBSTACLES

Purposeful restriction of work output has long been recognized as one of the indicators of economic insecurity on the part of industrial workers. Roethlisberger and Dickson (1943) have shown us in the Hawthorne study how the "rate buster" (worker who overproduced) was subjected to strong social pressure from his fellows to bring his production down to the group norm. The belief behind this practice was that overproduction would allow the company to introduce pay cuts.

Many management groups have attempted to deny the reality of this economic insecurity by fastening the blame for restriction of output upon labor unions. Mathewson (1931), however, has shown in an extensive study that restriction of output is equally common in plants in which none of the employees belong to unions.

If management will look a little more closely at this practice of restricting individual work output, they will see merely a repetition on a smaller scale of what the company does when it cuts down on production for fear of "flooding the market." Economic insecurity in a competitive economy cannot easily be ignored. It will be important then, in this connection, to study carefully the results of early attempts at introducing annual wage plans into the industrial economy.

Technological unemployment, no matter how temporary, is still a real threat to the worker. The studies of Nyman (1934) and Selekman (1945) have shown how disturbing this fear can be, how it is reflected in lowered productivity, and how it can reach a point at which any change in technical procedures may result in anxiety and resistance. Unless the worker can participate in the changes so that his sense of personal worth is not infringed upon by the introduction of labor-saving devices and unless he can somehow share in the economic savings afforded by the new equipment without any loss to his job security, this problem is likely to continue to be a serious source of friction between management and labor.

Another technological problem is represented in the feelings of social isolation and personal inadequacy found among workers engaged in belt-line production. Mayo and Lombard (1944) have shown how the breakdown of the work line into small, semi-independent teams of workers increased production and decreased absenteeism and labor turnover markedly in several aircraft plants on the West Coast. Chaplin's film satire (*Modern Times*) may have done much to extend recognition of this problem. Although heavy investment in capital goods has prevented the automobile industry from making significant modifications in its belt-line procedure, a great number of organizations in other industries seem to be aware of the dangers involved, as well as the general psychological drawbacks to superefficient time-and-motion procedures.

The sheer size of modern industrial organization represents not only a problem in communication, as discussed in a previous section, but a problem of individual impersonality. The large organization often seems cold and formal to the individual. He feels small and powerless, just another "cog in the machine." Part of this feeling is due to the fact that the efficiently run, large organization must reduce its individual personnel problems to over-all types and set up rules and regulations that, as Merton (1945) points out in his analysis of bureaucratic structure, tend to ignore individual variation to a great extent. The sense of impersonality induced by such factors could very likely be significantly lessened by arranging for employees to participate somehow in the setting up of rules and regulations. Although this type of problem is mainly characteristic of only the unusually large organizations, the problems of communication already discussed seem to be fairly prevalent in organizations of all sizes beyond the distinctly miniature.

The last, and perhaps the most crucial, problem to be discussed is concerned with the difficulty in modifying the unrealistic attitudes and inefficient supervisory habits of managers whose cooperation is needed to insure the success of new human relations programs. The same problem exists with first-line supervisors who carry out these programs at the point that counts most. We might say that good leadership in solving morale problems and in supervising people requires effective attitudes, adequate skills, and sufficient information.

Although personnel selection techniques are rapidly approaching the stage at which the effectiveness of supervisors may be successfully predicted at the time of hiring, the more important problem facing industry today is the retraining or re-educating of supervisors and managers already employed. Unfortunately, the most effective manner in which to carry on this training has not been determined.

Experience indicates that industry's partial failure to maintain good lines of communication between the top and bottom of the industrial ladder has resulted in many unrealistic, stereotyped attitudes on the part of both groups. Experience with management groups has shown that lecture and literature, singly or in combination, are not adequate. One promising technique, however, is the informal discussion group in which the members are made to feel free to participate vigorously, "let their hair way down," and examine the background of their own beliefs. This means that the group as a whole must be won over to the procedure; no one member is likely to examine his beliefs without being defensive unless he feels that the others are all committed to the same policy. French and Marrow (1945) reported one very impressive case in which the management's unrealistic attitudes toward female workers underwent a striking change during group discussion. An illuminating footnote to the report explained that a newly hired junior executive who held the old attitudes would not change his views when subject to pressure by the "reformed" managers, even though he was a minority of one; later on, however, when he was able to share in a group discussion with other managerial recruits, he was able to change his views.

At present a great number of widely located attempts are being made to study and evaluate this conference and discussion technique. At the moment, these procedures show great promise of becoming a major tool in the re-educating of supervisory personnel.

READING REFERENCES

Arensberg, C. M., and D. L. MacGregor, "Determination of Morale in an Industrial Company." *Applied Anthropology,* 1942, **1**, 12–34.

Drucker, P. F., "The Way to Industrial Peace." *Harper's Magazine,* December, 1946; January, 1947.

Hoslett, S. D., *Human Factors in Management.* Parkville, Missouri: Parks College Press, 1946.

Huszar, G. B., *Practical Applications of Democracy.* New York: Harper & Brothers, 1945.

Rennie, T. A. C., et al., "Toward Industrial Mental Health." *Mental Hygiene,* 1947, **1**, 66–88.

Roethlisberger, F. J., *Management and Morale.* Cambridge, Mass.: Harvard University Press, 1947.

Watson, G., *Civilian Morale.* Boston: Houghton Mifflin Company, 1942.

Whitehead, T. N., *Leadership in a Free Society.* Cambridge, Mass.: Harvard University Press, 1936.

47. MILITARY MORALE PROCEDURES

FRANCIS KEPPEL

Dean, Faculty of Education, Harvard University. Formerly: Assistant Dean, Harvard College; Secretary of Joint Army and Navy Committee on Welfare and Recreation; Information and Education Division, US Army; Assistant to the Provost, Harvard University.

INTRODUCTION

Morale always has been recognized as essential to victory. Military commanders have been concerned with the attitudes of their troops since earliest recorded history, and some, like Napoleon, have rated morale as having a value of 3 to 1 compared with the material equipment of an army. The daily necessities of military life impress on any leader, at any time, the importance of coordinated and willing teamwork. The conditions in combat give to it the sanction of life or death. It is little wonder that since time immemorial maintaining the morale—the *ésprit de corps*—of his unit has been a major responsibility of the line officer.

In World War I, perhaps because it was the first major and prolonged war since the development of democratic institutions throughout Europe, the defections of armies became a matter of such concern to the governments that separate administration of morale activities was instituted. The disaster of the Italian army at Caporetto in 1917, which is usually assigned to low morale, was followed by the establishment of a special branch of the army general staff to direct morale work. Following the severe strain of three years of war on French territory, in 1917 the French developed an organization of special officers for the stimulation of the morale of the poilus.

In American history the story of morale work appropriately begins with Washington's armies. His efforts to obtain the help of "a small traveling press to accompany Headquarters . . . an ingenious man to accompany this press and be employed in writing for it might render it singularly beneficial" fell on unsympathetic ears in the Congress, but he found and used a one-man morale officer in Thomas Paine. Paine lived and marched with the troops. He was the first of the G.I. correspondents, and *The Crisis* was as well known to American soldiers more than a century and a half ago as were Bill Mauldin's cartoons to the G.I.'s in 1945.

President Lincoln showed great concern over the defeatist Copperhead propaganda sent to the armies of the North in the Civil War, and the role of the civilian press in the Spanish-American War has often been described. In America it is typical that the morale of the army is a matter of public concern, for the soldiers have not been hired Hessians in time of war; they have been average American citizens. The withering blasts of criticism with which Americans in or out of uniform often favor their professional soldiers, even after a successful war, are examples of a distrust of militarism that is apparently deeply rooted in the American character.

The analysis of morale in the American army and the action taken on the basis of such analyses are therefore of both special importance and unusual difficulty. It may be of particular interest to the student of applied psychology, for it compresses into short time periods and makes more obvious the basic problems faced by the users of all social science research in a democracy.

THE TWO WORLD WARS

WORLD WAR I

The American Army was largely unprepared for the analysis of troop morale at the start of World War I. The most pressing problems in the years preceding the declaration of war had been the social and moral conditions at the Mexican border. A partial result was that in the opening months of the war great emphasis was placed on the recreation and welfare of the soldiers; in fact, the tendency to equate these matters with morale in the military and public mind lasted until World War II.

The first effort to apply more vigorous analysis to the actual attitudes of troops, and the forces which influenced these attitudes, was led by Brigadier General (then Colonel) E. L. Munson. At Camp Greenleaf in 1918 an organized program to stimulate morale of new recruits was tried out in the detention camp and later expanded to the whole cantonment. The experiment was conducted by the psychologists of the School of Military Psychology at the camp. The purpose of the experiment was to establish a morale program molding like-mindedness in the convictions and activities of soldiers and included both a personal effort to guide the recruit through personal contacts and accurate information and a coordinated group effort combining recreational, religious, and educational activities. The aim was not only to adjust the individual in his first contacts with the army but to create an environment through which he could be adapted with relative ease to his later army experience.

Greenleaf was, however, but one of many camps, and its techniques were unique. Morale was usually thought of in terms of entertainment and welfare activities, which were managed by private civilian agencies under a government commission. Though thousands of commanders used the best techniques of applied psychology in leading their troops, they would perhaps have been surprised had they been told they were doing so. It was not until the end of the war, when the dissatisfaction of troops in Europe needed no sensitive instrument of measurement to show its importance, that a Morale Branch was established in the General Staff, both to coordinate the work of the civilian welfare agencies and to consider and take appropriate action on other factors that influenced the morale of the army. Some experimental studies were undertaken and other work was begun that might well have had important implications for the military establishment in the years that were to follow. The drastic curtailment of appropriations to the Department of the Army, however, inevitably struck hard at any unit so lately established, and its only important deposit was the book by the Chief of Branch, General Munson, *Management of Men*.

Though limited because inadequate and insufficient social research had been carried on in this field, General Munson was years ahead of his time, and laid a groundwork on which a structure of sound policy in leadership and personnel management might well have been built. His broad conception of the foundation of military morale was largely forgotten, however, by 1939, and when Mobilization Regulations 1–2 (October 6, 1930, Change 1) Mobilization Regulations 1–10 (June 9, 1938), and Mobilization Regulations 1–10 (October 28, 1939), were written, morale was defined largely in terms of physical welfare: food, leaves, medical care, and the like, and morale activities in terms of recreation and welfare work.

WORLD WAR II

In early 1940 the War Department established in the office of the Adjutant General a Morale Division. It included subdivisions to handle entertainment films and other types of recreation, the award of decorations, and a section entitled "Morale Publicity." Actually this last unit was concerned primarily with public relations rather than with the information of troops. Except for matters of detail, these *Morale* activities did not differ markedly from the program of World War I, but their sponsorship and execution was, for the first time, in the hands of the War Department. This was a most significant change, since it fixed on the Army direct responsibility for the morale of troops in wartime. As we have indicated above, in the Revolutionary War, the Spanish-American War, and World War I private organizations provided both the initiative and the funds for many activities dealing with the welfare of troops.

During the period between 1940 and 1943, the officers in charge of morale work made a basic change in War Department thinking. Three major factors lay behind this change. First was the impact on the Army of vast numbers of drafted troops. Almost all of these men had had no previous experience of the Army, were in no sense Army-minded, and indeed, some were cynical about their part in the war. Second was the fact that the Morale Office in Washington expanded rapidly to include civilian experts in or out of uniform who had had training in the social sciences. The general in charge, Frederick H. Osborn, combined considerable business experience with extensive studies in several areas in the social sciences. Third was the fact that there was made available to the Army, for the first time, objective analyses of the attitudes of the drafted army. These were the result of a newly created research branch under the leadership of Professor Samuel A. Stouffer. The first results of this research, though often and necessarily fragmentary, tended to show that the provision of recreation and welfare activities for soldiers was not a major factor in the creation or maintenance of high morale. There were indications that the absence of adequate facilities would tend to depress the morale, but that such factors as job assignment, the competence of noncommissioned and commissioned officers, and a sense of mission were of far greater importance.

Since the Morale Division (which in the years to follow changed its name several times and finally became the Information and Education Division) had no direct control over either personnel assignments or the nature of army training, it naturally turned its attention to the provision of materials for the information and education of the army. As a result, by the end of the war, this division had taken a major part in the establishment of the largest single program of publications, radio broadcasts, and film showings ever undertaken in the history of the United States. For example, *Yank* reached a worldwide circulation of 2½ million copies a week, and the Armed Forces Radio Service had established 177 stations of its own and a complex affiliation with the radio stations of foreign countries. News was broadcast almost every hour wherever there were Americans in uniform. The largest correspondence school in the world had enrolled its millionth student.

Similar efforts were made by every army in World War II with varying degrees of success. Many of the Army's programs were also used by the Navy. The unique aspect of the program of the American Army in determining policy, however, was its attempt to make use of rigorous research on soldier attitudes. Such studies covered a wide range. For example, there are listed below ten studies which the Director of the Division believed in 1945 to have had particular significance.

1. Point Score System
2. Attitude of infantrymen and other soldiers toward the Infantry
3. The effectiveness of physical conditioning programs in Basic Training
4. Studies for orientation programs for D-Day in the ETO (European Theatre of Operations)
5. Studies of the Army Air Forces in Europe, including material on men's attitudes toward numbers of missions, furloughs, and other topics
6. Attitudes of ETO soldiers toward direct redeployment to the Pacific
7. Fraternization
8. Development of psychoneurotic screening tests used since October, 1944, to aid psychiatrists in screening men called for induction

9. Statistical studies of casualties in MTO (Mediterranean Theatre of Operations) in relation to replacement needs

10. Special studies at the request of the Surgeon General (trench foot), the Quartermaster General (winter clothing), and other War Department agencies.

One should not conclude from the preceding remarks, however, that the attitude of troops was consulted on all appropriate matters of policy. Rather, research results were often used in the justification of policies already arrived at; this will scarcely come as a surprise to any student of applied psychology who has read the complaints of social scientists in many civilian activities. However, a detailed study of the procedure followed in working out the Point Score System of discharge used by both the Army and the Navy in 1945, would indicate some of the issues involved, not only in obtaining the raw information but in making predictions, on the basis of this information, of the probable attitudes of individuals under varying conditions, and finally in the general administrative problems of putting such a program into effect. Despite these problems many observers believed, after the demobilization had been completed, that the Point Score System, based largely on attitude research, was as effective and as fair a way of demobilizing an immense army and navy as could have been devised under the circumstances.

With this brief historical summary in mind it becomes clearer why the American Army by the end of World War II had abandoned the conception of a morale staff, as such. The factors affecting the attitudes and motivations of troops had been shown to be so varied and complex as to surpass the capacity of any one staff section or staff officer. The very word "morale" itself had disappeared from the official War Department organization charts. To an increasing extent, emphasis was placed on good leadership at all levels of command. The reports of the Research Branch of the Information and Education Division, which emphasized again and again the importance of such leadership in the attitudes of troops, tended therefore to reinforce a view long held by professional soldiers: morale was the commander's responsibility.

The Research Branch studies, however, suggested a qualification to this dictum: morale was indeed the commander's responsibility, but his task would be much easier if basic War Department policies were, as far as possible, developed with an eye to soldier attitudes. In World War II, for example, many officers believed that the policies governing rotation of personnel to the United States from overseas assignments, the replacement of front-line troops, and the methods of handling complaints could have been markedly improved if subjected in advance to rigorous analysis, by trained psychologists and sociologists. Commanding officers whose duty it was to enforce existing policies were sometimes hard put to make them acceptable or even understandable to their men.

ADMINISTRATION

When the complexities of developing and maintaining high morale are considered in the light of the developments of World War II, it is scarcely surprising that there should have been, as there was, confusion over the most effective administrative organization. Three patterns are possible in the modern army. Authority may rest (1) with the line organization, (2) with a central administrative office with field representatives, and (3) with a combination of both. Under line administration, the commanding officer is responsible with little help in policy, training, or materials from higher echelons. This allocation of authority was characteristic of armies of all nations up to the time of World War I. Under central administration, morale officers are appointed for army units or camps to work with line officers but with independent authority under the direction of a Chief Morale Officer. It was toward this scheme that the American Army was pointing when in 1918 the Morale Branch of the General Staff was established for "the initiation and administration of plans and measures to stimulate and maintain morale of troops" (G.O. No. 94, 1918). In later years the political commissars of the Russian army seemed to have somewhat similar authority and administrative position, except that their authority rested outside the army chain of command.

A combination of these two forms of line and central authority is the more common today and is the basis for the policies of the American Army. Special staff sections at War Department headquarters are responsible for over-all analyses of factors affecting morale, for providing active programs of entertainment and recreation (Special Service Division), and for providing opportunities for an understanding of the reasons for military service, as well as opportunities for personal advancement while in uniform (Information and Education Division).

Staff officers assigned to field units are responsible to their commanding officers for such varied tasks as the proper coordination of military and nonmilitary agencies helping to entertain or provide welfare services for troops, the management of athletic events, the establishment of educational programs, the publication of newspapers, the organization of discussion groups, and the like. Where such an officer has earned the confidence of his commander, he may even extend his activities to the investigation and correction of undesirable conditions of any kind affecting morale. In World War II, however, it was rare that an officer achieved such a position, often because his responsibilities were so varied and onerous, and his time and energy so limited, that such investigations were almost impossible.

When, in the early years of World War II, the American Army tried to staff its regiments and other units with officers qualified for such broad morale responsibilities both of analysis and administration, it became evident that qualified personnel were almost impossible to obtain in large numbers. The trained psychologist was rarely both an educational administrator and a manager of motion picture theatres. The tendency in later years was toward the assignment of experts to these administrative tasks, and the training of all officers in the fundamentals of leadership and personnel management. Postwar revision of officer-training programs often included plans for such training.

The scientific analysis of attitudes, and the predictions of changes of attitudes, in the Army were found to be such complex and delicate tasks that it was necessary to centralize all such studies in the Washington office (with overseas staff when necessary), under the direction of qualified experts in psychology, sociology, and related disciplines. The danger of amateur efforts was great; indeed, hasty misinterpretations of published reports of the central office may occasionally have resulted in more harm than good.

The administrative dilemma is clear: since every major department or service in an organization so vast as a modern army or navy develops plans and policies that affect the motivations and attitudes of personnel, it would theoretically be desirable to have a qualified research unit immediately available to provide information when these plans and policies are made. Yet personnel and budgetary limitations make such a development impossible, and experience shows that amateur effort is often worse than nothing. A compromise effort was made in World War II to keep the Army informed in general of the attitudes and the morale of troops on a wide variety of topics by a monthly publication, *What the Soldier Thinks,* by wide distribution of special reports, and by encouraging staff sections in Washington and in overseas headquarters to consult the research staff on any problems on which they thought it might be of help. As a result of these studies and publications a considerable body of information on the background, education, aspirations, and attitudes of the soldier was widely distributed and doubtless had considerable influence on the day-by-day decisions of officers and noncommissioned officers. Much of this information has permanent value for sociological and psychological research and is now being studied intensively by a committee of The Social Science Research Council under a grant from the Carnegie Corporation of New York.

MORALE IN THE ARMY OF A DEMOCRACY

Discipline in a citizen army presents special problems. Wise leaders in the American Army, in time of war, have sought to achieve a combination of discipline and acceptance of authority on the one hand and of independent thinking and initiative on the other. These two almost contradictory aims might be described as mass discipline and self-discipline. In the former the soldier is trained to do as he is told, and in the latter he lives as a member of a social unit in which individual initiative for his own benefit and for the benefit of the group is encouraged. In a sense, this latter kind of discipline is a continuation of the soldier's role as a citizen out of uniform. In this capacity the soldier is used to freedom of speech, to the chance for self-improvement and advancement, and to the right to know why he is assigned to various kinds of tasks. It is no wonder that conflicts arose between these two sides of the soldier's life, which were reflected from time to time, for example, in the Army press, the management of discussion groups, and in the soldiers' letters home.

The existence of this dichotomy in a democratic army lies at the center of morale programs and argues for the need of a continuing research program not only into the social structure of the army and its attitudes but also into the technique of personnel management most effective in a military environment. Certain pilot studies made during World Wars I and II have shown the path. The peculiar advantages of the military organization, in which control studies under carefully ordered conditions are far more possible than in civilian life, should not be forgotten; the results of such research may well be of importance as well in civilian fields.

SUGGESTIONS FROM OTHER MILITARY MORALE ORGANIZATIONS

In contrast to the problem presented by the anti-militarist soldier of the American army, accustomed to free institutions, is that of the German army. Total education in Germany was mental, spiritual, and physical preparation for total war and had its climax in military education. The programs of morale stimulation were essentially the same for soldier and civilian.

German programs in World Wars I and II were not unlike except in detail and extent of application. Propaganda consisted of heroics, lectures, literature, and reports of glorious victories, sinkings of enemy ships, enemy weaknesses, vast preparations for the final offensive, and promises of glory, final victory, and the new order. In World War II this propaganda was necessarily spread wider to civilians, and the control of counter propaganda was greater because in war the civilian was a partner with the soldier to a greater degree. Morale discipline was strengthened through emotion: for example, love and sacrifice for *Vaterland, Gott, und Kaiser* during World War I and "One Reich, One People, One Fuehrer" during World War II. Propaganda of impregnability supported civilian and army morale; for example, the Hindenburg Line in 1918 and the Siegfried Line in 1939. To develop proper morale attitudes, discussions by soldiers were encouraged. Trained lecturers talked to German soldiers during World War I upon such subjects as "The Calamity of Defeat." In the training of the Nazi army, soldiers gathered together every day for explanation by their officers of current political events so that all might have immediate participation in them. Officers told of experience in previous combat in speeches called "ten-minute front spirit." Physical prowess was encouraged. The final morale technique of fear for civilians and army was entrusted to Himmler.

SPECIAL PROGRAMS

Conscription has always brought with it special morale problems, as was indicated in the US Army in 1917 and 1940. In Nazi Germany the special program for drafted recruits in 1935 consisted of a welcome to camp with fanfare, for conscription was an honor; the recruit's personality was studied in psychological examinations or interviews; and he was assigned a room-comrade of similar social status. Placement of congenial persons interested in similar arts and crafts in the same regiment is reported to have had astonishing morale effects.

German psychologists developed remedial programs for homesickness, suicidal tendencies, sex desire, homosexuality, individualism, and intellectualism. Homesickness therapy in the German drafted army consisted of frequent furloughs during early adjustment to army life. Numbers of furloughs were reduced as habituation to army life took place. Experiments were conducted in Germany, in preparation for World War II to control fears of isolation, protracted bombardment, superstition, surprise, suspense, gas, and panic among soldiers.

SOLDIER MORALE AND THE HOME FRONT

The close relation between soldier and civilian attitudes, emphasized by the Germans, necessarily was a considerable factor also in the American Army. In a sense, the great majority of drafted soldiers and sailors never stopped being civilians. The constant contact with American life, even when overseas, through newspapers, magazines, films, and radio, and through letters and visits home, meant that the soldier was sensitive to the attitudes and ideas of the civilian public. The student of army morale therefore must always carry in mind the influences brought to bear simultaneously on the civilian public.

Detailed descriptions of programs of psychological warfare are to be found elsewhere in this volume. It is enough to note here that the problems of morale in any army are inextricably linked with the climate of thought and the traditions of the nation which that army serves, and with the efforts of the nation itself, and of the enemy, to influence the thinking of its people.

READING REFERENCES

Bartlett, F. C., *Psychology and the Soldier.* Cambridge, England: Cambridge University Press, 1927.

Dollard, J., *Fear in Battle,* Washington, D.C.: The Infantry Journal, 1944.

Houle, C. O., et al., *The Armed Services and Adult Education.* Washington, D.C.: American Council on Education, 1947.

Munson, E. L., *The Management of Men.* New York: Henry Holt and Company, Inc., 1921.

War Department, *Information and Education Officer.* Washington, D.C.: Government Printing Office, Technical Manual No. 28–210, 1945.

48. COLOR VISION

DEAN FARNSWORTH

Lt. Commander, USNR; Head of Color Vision Selection and Research, Medical Research Laboratory, US Naval Submarine Base, New London, Connecticut; Color and Lighting Consultant; Chairman of the Technical Committee on Color and Illumination, Illuminating Engineering Society. Formerly: Technical Director of the University Theatre, New York University; Scene Designer and Stage Lighting Consultant.

Opinions or conclusions contained in this article are those of the author. They are not to be construed as necessarily reflecting the views or endorsement of the Navy Department.

Normal color vision is required for certain occupations in most industries, but it is essential in paint and printing ink manufacture, in chemistry, horticulture, photography and photoengraving, in the manufacture of ceramics, textiles, and paper, and in illuminating and electrical engineering, as well as in art and decoration. Employees in other businesses have made costly mistakes in the purchasing and sales divisions because they lacked normal color vision. Any person thus handicapped is a hazard in a manufacturing operation which uses color codes. Since defective color vision occurs in about 9 per cent of males and nearly 1 per cent of females in the United States, certain steps are necessary to achieve efficient placement and utilization of workers.

The term "colorblindness" is misleading, but it has been established by long usage. It does not mean any type of "blindness," but indicates an abnormal or deficient system of color perception. Literature on the facts of color blindness has been compiled by Judd (1943). Of chief industrial significance are the facts that colorblindness (1) is inherited, (2) is not curable, (3) occurs in degrees of severity which are detectable by certain tests when properly administered.

Defective color vision is inherited, usually by sons of women who are normal but whose fathers are colorblind. There is no reason to believe that color vision changes materially with age. Garth (1933) found that the incidence of defect varies widely with racial or cultural patterns; it ranges from about 2 per cent in male Indians (Clements, 1930) to more than 8 per cent in European males (Waaler, 1927; Von Planta, 1928).

Color deficiency is not curable, and has not yielded to any type of treatment, including the use of vitamins, diet, or electrical therapy. It is not correctable by colored filters, though lights or filters can be used as aids to color identification in certain specific situations. Such crutches do not, however, increase the color vision of the individual.

Industrially, the most important fact is that there are degrees of color deficiency ranging from a very slight deviation from normal to complete dichromacy (Chapanis, 1944). Dichromacy is the extreme condition in which all stimuli appear as yellowish, bluish, or greyish. The following chart gives a rough indication of the colors which are confusable by color defectives, though not to the same extent for all defectives. A more accurate analysis will be found in the color plates of Woodworth & Marquis (1947). It is important that the different degrees should not be lumped together, because men of moderate deficiency are capable of safely performing tasks which could not be trusted to dichromats.

A Chart of Color Relations as Seen by Color Defectives

Colors in the same line appear to be similar to color defectives, but any color in one line can be distinguished from a color in another line. For instance, pink can be distinguished from grass-green, because the pink may appear blue and the green may appear yellow.

grass-green chartreuse yellow amber orange-red
green olive tan brown red
dark green grey crimson
aqua powder pink
blue slate magenta
violet purple

There is also distinction by type of colorblindness which is of importance to science but not to industry, since the reactions of the two principal types to common industrial tasks are substantially the same.

And certain rare types may be neglected because of their infrequency. It is claimed that colorblindness may sometimes be acquired by disease or injury to the optic tract, but such disturbances are accompanied by more serious visual defects. However, the presence of congenital colorblindness does not imply any association with imperfect vision, outside of an unimportant reduction in the brightness and acuity functions. There is no reliable evidence that alcohol, tobacco, or fatigue impairs the color sense.

The various degrees of color deficiency can be detected by a series of tests which, however, must be properly administered in order to be valid. The commonest type of test consists of confusion charts, also known as polychromatic or pseudo-isochromatic plates, of which the Stilling and Ishihara, and the reprints published by American Optical Company are the best known. These plates have a spurious ease of administration; one half of the test construction consists of proper lighting, an element which is frequently ignored, thus invalidating test results. The colors of such plates are prepared for use under average daylight, i.e., illuminant C at approximately 6500° K. But natural daylight is extremely variable, and Hardy, Rand and Rittler (1946) have shown that unless a Macbeth daylight lamp is regularly used in the examination, no reliability can be placed upon the results.

When properly administered, confusion chart tests are failed by from 8 to 9 per cent of an average American male population. They are failed by the borderline, the color weak, and the colorblind alike. Personnel who fail confusion charts cannot be used as color inspectors or graders for cotton, butter, hay, textiles, plastics, or dyestuffs; or in colorimetry of oil, beer, or medical specimens. However, at least half of the 8 to 9 per cent who fail confusion charts are but partially deficient and can be used in ordinary industrial services involving the use of near-by, clearly distinguishable, code colors such as stop and go lights and the safety paint code devised by Birren (1945). In order to fill the need for a less stringent test, Farnsworth (1943) in cooperation with the Munsell Color Company and members of the Inter-Society Color Council, designed an intermediate test (The Dichotomous Test) which fails only the severely defective 5 to 6 per cent of an average American male population. Personnel who fail the Dichotomous Test cannot, without mishap, perform as filing clerks when the files are color-keyed; cannot safely distinguish color bandings on piping systems; cannot wire coded electrical parts without error. They would be unreliable handlers of women's wear, stage lighting, or home furnishings. But those who pass the test, even though they fail confusion charts, are acceptable for ordinary tasks.

Fairly good color vision is required by motor vehicle drivers, and a higher degree of color efficiency for Navy, Coast Guard, Maritime, and Aviation personnel, and by railroad employees. Numerous lantern tests on the order of the Edridge-Green (1911) have been designed to simulate their principal color task, which is the sure recognition of lights at a distance, often obscured by rain, smoke, or fog. Lanterns are not reliable unless the optics, the mechanics, and the operation are standardized. Such tests should fail 6 to 7 per cent of males, permitting only the 2 per cent borderline to pass.

The chief danger to industry is the surprising number of men who reach maturity without discovering that they are color defective. Many such do not realize the extent of their defect until they have pulled a severe "boner" or caused an accident as the result of it. Color defective men can be safely entrusted with many specialized jobs if they realize their defect, and learn that they must never make a color judgment. Vocational and pre-employment tests should be adequate to eliminate this hazard. To prevent the dissipation of time and money in training individuals for a profession or occupation for which their color vision is not adequate, Pitt (1942) emphasizes the need for intensive testing of school children.

As job analyses become available for more and more occupations, it will become easier to specify the degree of efficiency of color vision which should be demanded. While industry is primarily concerned with the elimination of those who are unfit for specific ordinary occupations, there are specialized jobs which require extremely acute vision. Superior color discrimination must not be thought of as the converse of colorblindness. Men with perfectly normal vision may or may not have good color discrimination. This is partially the result of natural aptitude and partially of training. Sometimes an experienced, well trained partially colorblind worker will excel a low discrimination normal, just as an experienced one-handed carpenter may be more capable than his two-handed helper. Only two tests have been developed in this field, the ISCC Color Aptitude for testing superior discrimination (Dimmick, 1943), and the F-M 100-Hue for the below-to-above average range of discrimination (Farnsworth, 1943).

SUMMARY

Modern industries are makng increasing use of color in their operations, but it is unnecessary and economically wasteful to stigmatize workers of all degrees of color anomaly as color blind, and unfit for industrial use. Placement of personnel should be preceded by an understanding of the degree of color perception required in the positions to be filled; when this is known, the appropriate tests for applicants should be administered under standard conditions.

No hope for improvement of defective color vision can be expected from any treatment, though education tends to improve somewhat the discrimination of both normal and color defective. Vocational guidance should acquaint a man with the extent of his deficiency, and prepare him to avoid occupations, or the particular phases of occupations, for which he is unfitted.

READING REFERENCES

Farnsworth, Dean, "Industrial Tests for Color Vision." Transactions American Academy of Ophthalmology and Otolaryngology, November-December, 1948, 156–157.

Judd, D. B., "Facts of Color-Blindness." *Journal of the Optical Society of America,* 1943, 33, 294–307.

The Physical Society, Color Group, *Report on Defective Colour Vision in Industry.* London: The Society, 1946.

Pitt, F. H. G., "Reports of the Committee upon the Physiology of Vision: XIV. Characteristics of Dichromatic Vision. With an Appendix on Anomalous Trichromatic Vision." *Special Report Series, Medical Research Council,* 1935, Ser. 200.

Wright, W. D., *Researches on Normal and Defective Color Vision.* St. Louis: The C. V. Mosby Company, 1947.

49. BIO–MECHANICS

JACK W. DUNLAP, PH.D.

President, Dunlap, and Associates, Inc., New York; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology. Formerly: Commander, USNR and Officer in Charge Aviation Free Gunnery Training and Research Unit; Director of Research, Committee on Selection and Training of Aircraft Pilots, National Research Council; Associate Professor of Educational Psychology, University of Rochester and Fordham University; President of Psychometric Society, New York State Association of Psychology, Division of Consulting Psychology, American Psychological Association, and Division of Psychology, New York Academy of Science; Editor, *Journal of Educational Psychology* and *Personnel Journal.*

Bio-mechanics is the application of scientific principles and data to the articulation of men and machines in order to enhance their efficiency and the job satisfaction of the operator. An important but narrow aspect of bio-mechanics relates to the design of equipment in terms of the limitations and abilities of the operator. The broader scope of bio-mechanics includes a study of the optimal work stations, the physical and interpersonal environment and, indeed, any set of conditions which may affect the productive capacity of men and machines as a working unit. Other terms currently applied to describe essentially the same field of study are: human engineering, psychological problems in equipment design, psychotechnology, applied psychophysics, psychoacoustics, psychophysical systems research, and systems research. The number of these titles are symbolic of the youth of this particular field of study. Minor differences in emphasis may be found among them, as, for example, in the field of psychophysical systems research a greater emphasis is placed on the problem of the organization and operation of numbers of men and machines with a common mission. A 4-engine bomber with its aircrew has been called an "aircraft system." Each such system is studied to improve the efficiency with which its mission may be accomplished.

Interest in bio-mechanics is probably as old as the first tool fashioned by primitive man. Undoubtedly, the form of the early hammer was not determined entirely fortuitously. The length and diameter of the handle and weight of the head were influenced probably by an intuitive recognition of the limitations in the size of the hand, in the strength of the forearm, and in the precision of controlling a long handle. While interest in bio-mechanical problems is ancient, a precise experimental approach to the interrelation of man and machines is relatively new. It is not too important to indicate when the first bio-mechanical study was performed, for bio-mechanics has so many different roots that such a determination would at best be difficult. The term "bio-mechanics" has appeared in the literature of psychology, engineering, physiology and medicine. It is sufficient for our purposes to know that the research activities of many psychologists during World War II provided the impetus for the recognition of bio-mechanics as a maturing, distinct, applied psychological science.

One of the distinguishing features of World War II was its highly technological character. Many different kinds of specialists sought to devise more efficient machines for seeking out the enemy and destroying him. Many of these devices embodied highly desirable technical features but could not be put to practical use because they were too complex for the average man to operate. It was not long before a small number of psychologists were asked to apply their techniques and energies to the problem of adapting equipment to the operator rather than to selecting men to fit the equipment. Their first efforts produced such gratifying results that they were encouraged by the authorities to attack other problems which had but a secondary connection with machines. In meeting these problems, they developed an experimental literature which has wide applicability not only to a variety of military but also to industrial problems. A series of examples has been selected to demonstrate the application of bio-mechanical data and approach to problems in the military and industrial fields.

BIO-MECHANICS APPLIED TO MILITARY PROBLEMS

COMPARISON OF EQUIPMENT

The operation of complicated military equipment usually requires considerable training on the part of the operator in order to realize its maximum potential. Tracking a moving aerial target accurately requires a high degree of skill. The aerial gunner must simultaneously estimate the speed, distance, and angle of flight of the enemy craft and combine these with similar data from his own plane to make a judgment on how to aim his guns. Such training, were it restricted to practice in the air, would be extremely costly in planes and flying and maintenance personnel. Special training devices of high transfer of training value, even if their cost is great, can easily represent a substantial saving in matériel and personnel.

Early in the war some officers in the Navy had to choose between two gunnery training devices, designated as *X* and *Y*; *X* was a relatively simple, double projection device, *Y* was more realistic but much more complicated. *Y* cost fifty times as much as *X*, required a crew of eight to operate, and was difficult to maintain. Before a final decision was made a psychologist was requested to evaluate the two pieces of equipment. To make the problem more difficult, no combat criteria were available to serve as a basis for determining the most effective instrument. Two groups of 150 men each were selected and matched carefully for experience, general intelligence, and mechanical aptitude. One group was given 8 hours of training on *X* and the other an equal amount of time on *Y*. At the end of the 8 hours the subjects were scored for proficiency on the instruments on which they had been trained. Each group was then transferred to the other machine, given a one-hour indoctrination course in its operation, and a score obtained on that device. Both groups proved equally efficient on the two trainers. It would appear, therefore, that a common core of learning was available with either of the two trainers. It was on the basis of these results that the simpler and less costly equipment was adopted by the Navy and later by the Army. During the war psychologists were frequently consulted in order to make similar decisions on the relative merits of competing equipment.

DISPLAY OF SOURCES OF INFORMATION

There are countless pieces of equipment used by the services which provide necessary information to a combatant or a combat team. Often, this information must be obtained, interpreted, and applied quickly and accurately. Examples of such devices are range finders, radar scopes, periscopes, and the instrument displays in aircraft, automobiles, tanks, submarines, and surface vessels.

The instrument panel of the modern airplane presents an extremely complicated array of information which must be perceived and assimilated in a very short time. The magnitude of this task may be appreciated from the fact that cockpits sometimes contain as many as 350 switches, dials, and gauges for the purpose of information and control. The dial may be located differently in different planes and the grouping of dials may be different from plane to plane. Clearly, the legibility of the dials and discriminability of the switches and controls are important problems. A dial contains a hand, a face, numbers, and other markings. The numbers may increase clockwise or counter-clockwise; the zero mark may be placed at different points on the circumference. Any or all of these factors may lead to confusion on the part of the operator. It was necessary to systematically alter each of these variables to secure information for redesigning dials.

It was demonstrated that accuracy in reading a dial is not enhanced by providing very fine subdivisions (Loucks, 1944). With exposures of three-quarters of a second, comparable to the duration of a pilot's glance, the dial with the fewest markings, illustrated in Figure 1, gave the smallest percentage of error. Another interesting finding was that the numbering of the smaller subdivisions tended to *decrease* the accuracy with which an instrument could be read during a short exposure. Within limits, the reduction in the width of the pointer did not improve accuracy and apparently the starting point of the dial did not influence legibility.

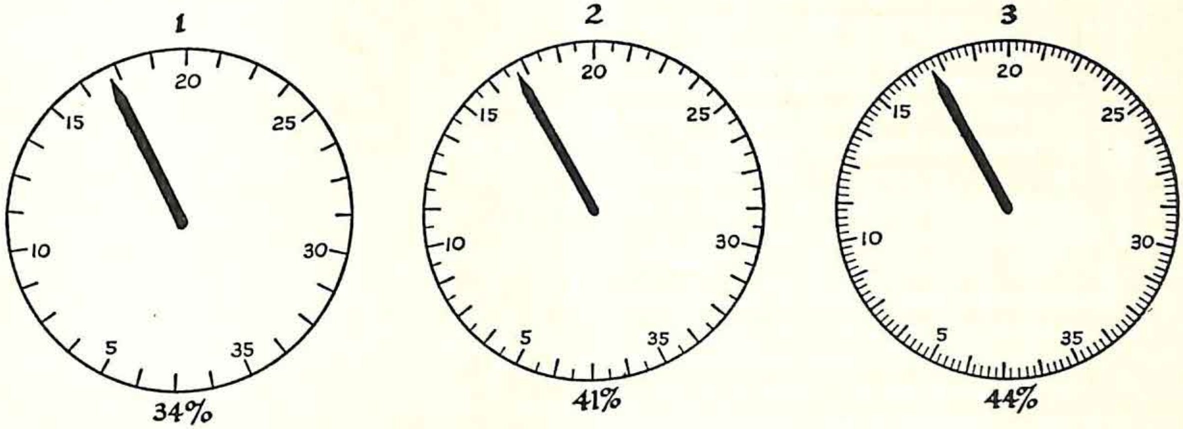


Figure 1 (VI–49): *Various Types of Dial Faces*

Accuracy in reading a dial is not enhanced by providing fine subdivisions. Figures under each dial indicate percentage of error in reading the dial with an exposure of ¾ second (Loucks, 1944).

The problems of the location, grouping, and even reduction of the number of instruments on the panel are now undergoing intensive study. It is generally agreed that the standardization of the location of instruments and controls in planes of different manufacture would contribute greatly to a reduction in accidents due to pilot error.

THE PROBLEM OF CONTROL.

Achieving accuracy and ease of operation of controls offer problems which correspond to those encountered in the brief survey of the display of information. When movement of a part of a device is required, what shall be the nature of the movement in order to achieve the *highest level of precision* with the greatest speed and least fatigue? Sometimes all these goals are equally desirable, and in some problems one or two are given priority over the remainder. Operator motion can be rotary by means of a knob or crank, clockwise or counter-clockwise; it can be of a lever type and the direction and degree related to the effect in different ways; the motion can be at different levels; it may involve the preferred or nonpreferred hand; it may be effected in the horizontal, vertical, or intermediate plane; the problem may require that the motion be made in unusual body positions. These are but some of the possible variations in applying motion to a device. The biomechanic analyzes the missions of the machine and their priority and evaluates and applies known experimental data to the machine being developed in accordance with these criteria. In the absence of known data, experiments are performed comparing the efficacy of different types of motions.

In one study (Grether, 1946) the relative accuracy of the hands and feet in carrying out the same control task was determined in an apparatus resembling a cockpit. A further comparison was made of the precision of lateral and fore-and-aft movements of a wheel and a stick-type control. The subject's task was to keep the pointer of an aircraft dial stationary. The pointer was made to wander in an erratic fashion and the subject was required to keep it centered by one of the conventional flight controls. It was found that the use of the rudders by the feet was less accurate than the use of the stick or wheel by the hands. No difference in precision was found between the stick and the wheel for aileron or elevator movements, but it was discovered that most pilots could make fore-and-aft or elevator control movements more accurately than lateral or aileron control movements.

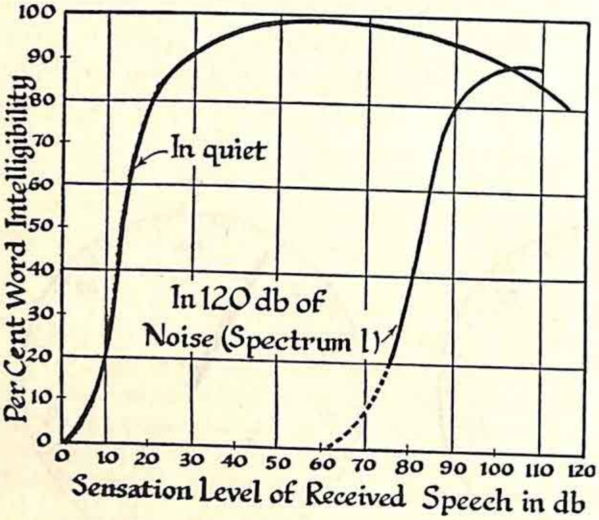


Figure 2 (VI–49): *A Comparison of Intensity Levels of Received Speech Necessary to Equate Intelligibility under Background Conditions of 120 db of Noise and of Quiet* (Egan, 1944)

In another study (Johnson, 1946), a comparison was made of the effectiveness of a knob-type versus a pistol-type handgrip in tracking and framing a moving target with a particular pedestal sight. It was found that the operator could remain on a target 25 per cent more of the time with a pistol-type grip. As Fitts (1947) has pointed out, a comparable gain in efficiency through the redesign of the mechanical or electrical components of the sight or computer could be achieved only at considerable cost.

COMMUNICATION PROBLEMS

The coordination of teams of personnel engaged in combat is a serious problem. Some form of speech communication is usually necessary and generally under conditions which are far from ideal. Consider the intercommunication necessary among members of a bomber crew. Not only is intercommunication required to coordinate the fire power of the bomber on enemy craft but in the event of serious damage the pilot must notify the rest of his crew to bail out. In numerous instances during the last war crew members bailed out unnecessarily because they failed to understand instructions from the cockpit. Failure of comprehension on other occasions was commonplace under conditions of noise developed by the engines, guns, flak, and propellers, in addition to such other factors as cold, anoxia, excitement, fatigue, and poor adjustment of headphones. Even more commonplace and dangerous were failures of communications between pilot and control tower in bad weather when the static was high and visibility near zero. These factors made it imperative to subject the problems of speech intelligibility to intense study. More than thirty factors have been found which influence intelligibility. For example, it has been determined that the intelligibility of speech is not so much a function of its intensity as it is of its relationship to the intensity of background noise. In another study the relationship between speech intensity and intelligibility with a background noise of 120 decibels was examined (Egan, 1944). These data are presented in Figure 2 together with data obtained against a background condition of quiet. Such knowledge is important to the design of headphones and of amplification circuits which can produce desirable intensity output for combat conditions.

Still another example of improvement in communication through bio-mechanical study is provided by research in "peak clipping." Most speech amplification systems have a nonlinear relationship between the input and output of energy. A high intensity signal input is apt to be disproportionately amplified and come out as a distorted or distorting signal. This distortion forces the operator to keep the average amplitude of speech sounds low. If a device called a "limiter" is placed in a receiver, it prevents amplification of any sound beyond a predetermined intensity level. Of course, this limitation provides a certain amount of distortion in itself, but simultaneously permits an increase in the intensity of the average speech sound with an even greater gain in intelligibility. The gain in speech intelligibility through the use of the limiter has been demonstrated empirically (Licklider, 1946).

VISUAL PROBLEMS

While audition is a very important sensory function, it would not be too difficult to defend the view that vision, particularly in wartime, is an even more important one. We have already considered certain visual problems in connection with the display of information in the cockpit of the airplane and elsewhere. Knowledge about visual functions obtained before the war could be applied immediately to a variety of military problems. For example, the retina contains two types of light sensitive elements, called rods and cones. The rods are very sensitive to light and insensitive to color, while the cones are sensitive to color and much less sensitive to light. The cones are around the center of the retina (fovea) while the rods are found mostly at the periphery and none are to be found in the fovea. This arrangement means that color perception is best when looking directly at an object, but under twilight vision, dark objects can be most easily detected out of the corner of the eye. This principle was applied in training military personnel to use off-center vision to detect enemy motion and to navigate in darkness.

Another fact well-known before the war is the difference in sensitivity of rods and cones to lights of different wave lengths. The rods are scarcely stimulated by a moderate amount of red light, whereas the cones are. This fact had many important military applications. Exposure to a bright light makes it difficult to see dimly lighted objects immediately. Night fighter pilots were able to preserve their visual sensitivity (rod vision) by wearing red goggles on the ground preparatory to taking off.

While flying at night, pilots alternate vision frequently between the cockpit and the sky when looking for other craft. Their visual sensitivity outside the cockpit depends, in part, on the intensity and wave length of the illumination inside. In view of the relative insensitivity of the rods to red illumination, it was found that red light in the cockpit would best preserve the light sensitivity of the eyes during night operations.

ENVIRONMENT

The environment in which a person can operate most effectively is a problem which has engaged the activities of many psychologists. A pilot's milieu can undergo dramatic changes in a short time. For example, many bombers were based in hot humid climates where temperatures of 110° F., and humidity exceeding 75 per cent were commonplace. While the bomber was being prepared for takeoff, exposed to the sun, the inside temperature often rose to above 150° F. In anticipation of high altitude flight the pilot and co-pilot would enter the craft in warm furlined jackets and trousers. Yet 30 minutes later, after climbing to altitude, the temperature might be as low as minus 30° F., with attendant discomfort from cold, perspiration, soaked garments. The pressure of the atmosphere changed from 760 mm. of mercury to possibly 175 mm. Physiological changes occurred symptomatic of oxygen deficiency. Pull-outs from steep dives often subjected the crew to great centrifugal (*g*) forces with attendant disruption of some of the normal physiological processes. When the direction of these *g* forces was from the head to the feet, members of the crew would "grey-out" or "black-out," or become unconscious depending on the individual and the magnitude of the force. When the plane would level off after its pull-out, some members of the crew would "red-out." These disturbances resulted from the unusual redistribution of the blood supply.

Early studies on aviation problems were largely confined to an interest in the critical limits of anoxia, extremes of temperature, of *g* forces, etc. "Critical" was then defined as physiologically harmful or even fatal. Subsequently it was recognized that a mistake of judgment or a delayed response could be just as fatal as acute anoxia, or extremes of temperature. "Critical" was then redefined to mean the level at which human performance deteriorated to an unacceptable degree. Simple and complex performances, and visual, auditory, and other sensory functions have been studied by psychologists under varying conditions of anoxia, temperature, *g* forces, vibration, noise, drugs, etc., and a substantial literature has developed. This work has influenced military specifications for safe operation of aircraft. For example, the lowest altitude at which an auxiliary oxygen supply was mandatory was determined to a large degree by the data collected by psychologists. Additional examples include the design of clothing to protect personnel from low temperature, the evaluation of air-conditioning equipment to guard critical personnel from the disturbing effects of high temperatures.

The foregoing brief account of the wartime effort of psychologists in solving bio-mechanical problems cannot begin to describe the magnitude of their contributions. It was intended merely to provide an idea of the scope of their interest in this area and of how experimental techniques were applied to military problems.

BIO-MECHANICS IN INDUSTRY

The principles of bio-mechanics can be applied to a wide variety of problems in the industrial field. Space does not permit the presentation of an extensive set of examples, but a few will suffice to indicate certain possibilities.

A PROBLEM IN RADIO DESIGN

A large manufacturer of radios questioned the advisability of automatic volume controls in radio receivers. The first question proposed to the psychologist was, "Do radio listeners need and want automatic volume control?" It was found that listeners did desire this feature, so the next question was, "How sensitive must the control unit be?" This question could be answered by experimental methods. Basically the problem was to determine the point at which individuals would adjust the volume control, regardless of the level at which the volume was initially set. Experiments provide that the volume could vary as much as 4 decibels before the listener would adjust the volume control. Thus it was possible for the psychologist to provide the engineer with definite limits of sensitivity, expressed in terms of the listeners' responses in designing the automatic control unit.

DESIGNING EQUIPMENT IN TERMS OF OPERATOR COMFORT AND EFFICIENCY

It is not an easy matter to find a piece of production equipment which will be familiar to the majority of students of psychology, but there is one with which many students come in contact, the IBM test scoring machine, which has been subjected to extensive developmental work both from the standpoint of operational efficiency and from that of the operator. In the earlier models the machine was studied from the standpoint of the operator. It was found to be too high, it required a footrest, it had sharp corners and projections which resulted in bumps and bruises, removing the answer sheets was inconvenient, the switches were poorly arranged, the illumination and angle of the dial face made reading difficult, and the feed slot was so placed as to require considerable twisting on the part of the operator. All of these inconveniences were noted and corrected, but experience has shown still other changes that can be made. (This phase of the developmental program had to be delayed, since the design was frozen for production during World War II.)

The operator sits before a flat console while manipulating the controls and feeds the answer sheets into the machine. The left hand manipulates four rotor type controls, two toggle switches, an operating switch and a starting switch. The right hand transports the answer sheets from a retainer, drops them into a scoring slot, records the score, and finally starts the motor to release the answer sheet.

The rotor switches require a 90 degree flexion of the wrist, which is inconvenient for many operators. Further, the switches are so arranged that the operator is forced to extend the left arm to nearly its full length in order to operate one of the switches. Both of these difficulties can be overcome by the use of push-pull lever-type switches, which are simpler to operate and can be located more conveniently for the operator. The right arm is nearly completely extended in transporting an answer sheet from the retainer to the scoring slot. Further, the load on the right hand can be reduced by controlling the starting switch with the foot. In a number of service installations, a foot control was employed, and it was reported that this change resulted in an increase in production of 25 per cent, and less operator fatigue. If this particular model, while in the mock-up stage, had been subjected to additional study in terms of the operator's limitations, it would have been more efficient as well as more acceptable to the operator.

THE USE OF CONTRAST IN SECURING ACCURACY

It happens not infrequently that an industrial problem can be solved by the simple application of a well-known psychological principle. For example, in the pharmaceutical business, an accurate count of tablets is a matter of prime concern for medical, legal, and other reasons. Paddles are generally employed in counting tablets. These paddles consist of an unpainted aluminum or wood surface in which a predetermined number of holes have been partially bored. The holes are slightly larger than the tablets which are to be counted. The operator shoves the paddle into a container of tablets, picks up a load of material and then shakes the paddle with a wrist motion until each hole contains a tablet; the excess is then poured off. It is then a matter of brief inspection to determine if all the holes contain tablets. This method of counting is simple and efficient but, it is not foolproof. Errors are made and these appear to be the fault of the operator. During routine production the operator occasionally fails to perceive that one of the holes does not contain a tablet, and consequently pours a short count into the bottle or other container. This perceptual error is due, in part, to the fact that the tablets usually are white and the paddle is also light in color. The precision of inspection was improved by the simple expedient of placing a spot of contrasting color, somewhat smaller than the diameter of the tablets, in the bottom of each hole. It was found that after this application of the principle of contrast that errors in counting were reduced appreciably.

A PSEUDO-SELECTION PROBLEM

It happens not infrequently that the psychologist is called on to help solve a problem which has been incorrectly formulated. Sometimes a cursory inspection of the elements of the situation will clearly indicate how the problem may be reformulated and attacked with satisfactory results to all concerned. Recently, the author was consulted on a problem in "selection." A certain industrial organization had an entire contract cancelled because five items were found defective in a lot of 100 gross. These items were worth but a fraction of a cent each, but the total contract ran to thousands of dollars. An executive of this organization, upon examining the problem decided, "What we need is a good selection system for inspectors." Therefore, a psychologist was called in to set up a selection system. The first question to be answered was, "What is to be selected?" So a thorough study had to be made to ascertain why certain defective items escaped the inspectors. Observation revealed that the items passed by the inspectors on a conveyor at the rate of 200 a minute. If an item was defective, the inspector removed it from the belt and placed it in a container at her side. When the container became heaped up she would bend down and level off the pile of material. In each case her eyes were away from the belt for several seconds. It was further noted, that many of the younger inspectors were apt to watch the young foreman as he moved about through the department. Although selection might help, particularly if elderly, unattractive men were the object of selection, the basic problem was to identify "attention distractors" and to develop means for reducing or eliminating them. Certain remedies are immediately clear, such as, rearrangement of the work station with regard to the disposal of defective items, training in the use of peripheral vision for handling defective items, to mention only a few.

THE APPLICATION OF THE RESULTS OF PUBLISHED RESEARCH TO THE WORK ENVIRONMENT

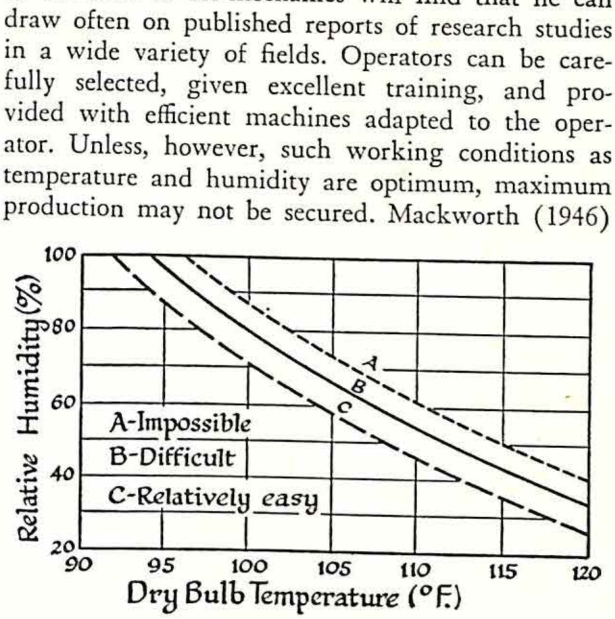


Figure 3 (VI–49): *Relationship between Relative Humidity and Dry Bulb Temperature in Determining Environments in Which Work Was Relatively Easy, Difficult, or Impossible for Acclimated Subjects* (Eichna, 1944)

The psychologist working on industrial problems in the field of bio-mechanics will find that he can draw often on published reports of research studies in a wide variety of fields. Operators can be carefully selected, given excellent training, and provided with efficient machines adapted to the operator. Unless, however, such working conditions as temperature and humidity are optimum, maximum production may not be secured. Mackworth (1946) studied the effect of various room temperatures on eleven wireless operators in terms of number of mistakes per hour. The number of mistakes per hour remained almost constant over the temperature range from 78° to 90° F., but the number of errors increased more than 100 per cent from 90° to 95° F., and at 97° F. the error rate was more than 400 per cent greater than at 78° F. Studies on humidity by the Armored Medical Research Laboratory at Fort Knox (Eichna, Ashe, Bean, and Shelley, 1944) conclusively demonstrated that it is much more difficult to work well if the temperature and humidity are both high. Examination of Figure 3 clearly shows that physical work becomes practically impossible when the temperature is 100° F. and the humidity is 86 per cent, while at 110° F., the same condition obtains with the humidity as low as 57 per cent. These two studies are supported by other investigations, but they are sufficient to demonstrate that atmospheric conditions which interfere with the normal body temperature, reduce physical efficiency. Extremes of heat and humidity can and do occur in industrial plants and must be taken into account if the workers are to produce efficiently. Recently in studying production in a southern mill, the workers were given a questionnaire in which they could express their attitudes toward working conditions, among other things. A very high percentage complained of the heat and humidity within the plant and strongly requested the plant be air-conditioned. At first, its management was inclined to treat the installation of such equipment as an unwarranted expense, but when the workers' wishes were supported with data as to the plant's temperature and humidity and with studies such as those mentioned above, the decision was reconsidered and air-conditioning has since been installed with gratifying results.

VARIABILITY IN THE PRODUCT OF AUTOMATIC MACHINES DUE TO THE OPERATOR

In the operation of automatic machines it is generally assumed that if the speed of the machine and supply of materials are kept constant, then the quality of the product will be constant. The modern manufacture of nylon hosiery entails the use of automatic machinery. The operator's function is to pull certain levers at certain times. When he will operate them is determined practically automatically as the machine completes each stage of its task. The operator cannot easily add to or subtract from the number of stitches or from the number of courses knit into a stocking. Superficially, it would seem that the operator could exert little influence on the finished product. The variations in length of foot and leg which do arise and which are reflected directly in differences in weight of individual stockings would seem to be attributable to variations in the nylon yarn and in machine performance.

To determine whether this assumption was justified, one investigator (Coakley, 1947) measured the weight of the nylon yarn, the weight of the finished product, and the variation in the performance of the machines. In weighing the unfinished and finished nylon, allowance was made for the presence of moisture and other foreign materials. Substantial variation (approximately 30 per cent) was found in stocking weight among different operators working on the same machine. Figure 4 shows the weight distribution of nylon hose produced by five consecutive shifts of three operators on one machine. The total number of hose equals 1,200. Weights are expressed as percentage deviation from total mean. All stockings were intended to be the same size. In addition, variations in mean stocking weight of as much as 15 per cent or more were found among different machines regardless of the operator. Such variation made for poor control of the stocking size; a stocking intended to be size 9½ would

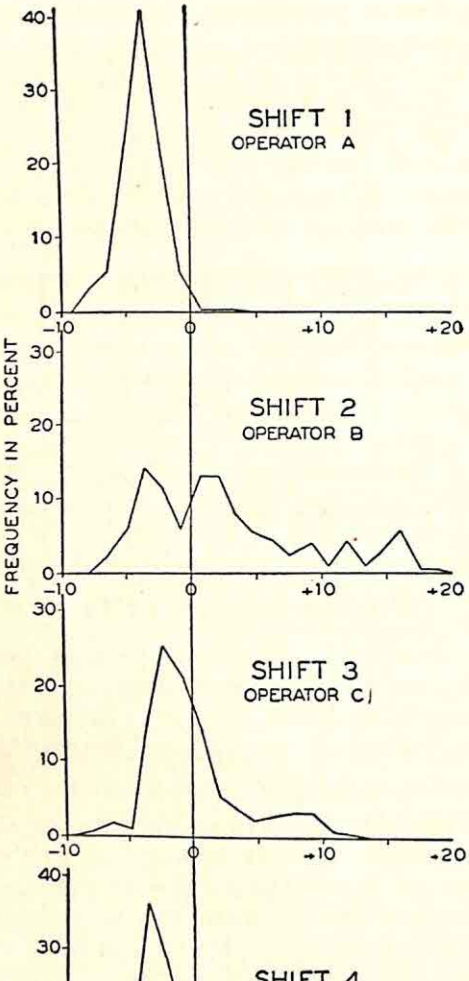


Figure 4 (VI–49): *Frequency Distribution of the Weights of Nylon Hose Produced by Three Operators on One Machine during Five Consecutive Work Shifts* (Coakley, 1947)

actually be size 10 or even 10½. Another undesirable feature of variability was the tendency to consume too much raw material. With the operator's contribution to variability established, the next step was to determine its causes and to design a training program to reduce it. This procedure saved at least 3 per cent of the yarn for an entire hosiery plant, in bad years the difference between profit and loss for many mills. Additional savings were effected through remedial work on the equipment. Here is an example of a psychologist's contribution to the improvement in quality control and conservation of raw material through the bio-mechanical study of the operators of highly automatic machines.

The foregoing examples are but a brief sample of the application of the principles and approach of bio-mechanics to the industrial area. How does this approach relate to others in the field of industrial psychology and in the field of engineering?

THE RELATIONSHIP OF BIO-MECHANICS TO OTHER ASPECTS OF INDUSTRIAL PSYCHOLOGY

Bio-mechanics in its broader aspects must support and in turn be supported by other phases of industrial psychology and the closely related field of time and motion study and labor relations. If production and operator satisfaction are to be optimal, the machine must be adapted to the man, the operator must be selected and trained, efficient work patterns developed, and the whole must be set in a satisfactory frame of human relations.

SELECTION

After the first World War many companies, as the result of the extensive use of tests in the war, organized personnel departments with the responsibility for testing and selecting employees. Whenever a plentiful supply of operators is available, it is practical to select those best qualified in terms of previous experience, or if experience is lacking, then in terms of certain characteristics of the task. This last objective requires a job analysis on the basis of which suitable tests may be selected or developed. It is advisable even when the machine has been completely engineered from the standpoint of the operator to select operators who will be the most proficient in using the machine. This topic has been treated at length in other sections so that it is necessary here merely to cite its interrelation with bio-mechanics.

TRAINING

Once an individual has been selected to operate a particular piece of equipment, he should be trained in the most efficient methods of operation. Instruction, rewards, and the kind of task involved, among other factors, determine the way in which an individual will learn, if at all. Yet, very often, industrial training consists in assigning the novice to an experienced operator and the novice is expected to learn by the "monkey see, monkey do" instructional method. This procedure often results in poorly trained workers, inefficient use of equipment, undue wastage of raw materials, and an undue number of breakdowns in equipment.

The development of a training course in industry requires considerable teaching skill. Effective training requires an analysis of the task into its fundamental components for emphasis without destroying its organization. The relationship between the various steps and processes in a chain of operations must be clarified and made dramatic so that learning may occur most efficiently. All available instructional aids, such as written material, charts, diagrams, other visual aids, oral instruction, and the actual operation of equipment, should be utilized. Often it will be found necessary to develop a standard vocabulary or patter, and in so doing one must keep in mind that few workers have a doctor's degree in their educational background.

TIME AND MOTION ENGINEERING

The efforts of time and motion engineers have resulted in substantial contributions to production, and through production, benefits to management and the employee. Unfortunately, in many cases, the results of motion analysis have been unwisely applied with the result that this work is connected with the odious term "speed up" in the minds of many workers.

The statement that "Man is a Machine" is frequently made and the casual acceptance of this statement by management, the labor relations expert, the psychologist, and the engineer has led directly to many of our current labor and production difficulties. It is true that a man can be used as a machine, to move raw materials, transmit communications, and to modify raw materials, to mention only a few ways in which man *resembles* a machine. Man's main difference from a machine lies in the fact that he is an emotional organism with a complex organization of interests, attitudes, drives, and desires for personal satisfactions. The failure to recognize and understand these attributes has resulted in conflict at enormous cost to both capital and labor as well as to the general public. It is imperative that these factors be studied with the same zeal and scientific detachment as has been devoted to the study of other aspects of the industrial situation.

Some of the activities of the bio-mechanic resemble those of the time and motion engineer, and others appear to be just the kind that would be found in industrial engineering. There is a very distinct difference in emphasis between the approach of the time and motion engineer and the human engineer. The time and motion engineer has traditionally considered the machine a constant and man a variable. Optimal movements were explored which were most suitable to continued operation on a particular machine by the average operator. While the time and motion engineer has given attention to the redesign of equipment, this activity has been peripheral to the main trend of his endeavors. Time and motion studies have been used by the human engineer to determine whether the number, speed, or character of the movements in the operation of a piece of equipment are excessive or maldistributed. Such a study is valuable to becoming acquainted with the demands that a piece of equipment make on the operator and is usually a preliminary to the modification of its design.

MORALE AND ATTITUDE STUDIES

In any work situation it is not sufficient to have the best equipment and the best possible operators unless the workers are happy and contented. The importance of morale has long been recognized but it is only recently that tools have been available for measuring attitudes toward various aspects of the work situation. The use of such techniques has spread rapidly since World War II, and proved to be of great value in identifying sources of dissatisfaction on the part of the worker. Many factors which appear trivial to the busy executive are important to the worker. For example, allowing a complaint by a worker to lie on an executive's desk two weeks may be misconstrued by the worker as an attitude of indifference. Or if a worker is suspended for 2 weeks, and through the press of other duties his case is not immediately reviewed, the delay is interpreted as disregard of the worker. Delay in reviewing a suspension is of the utmost importance to the individual concerned, though it may be inconsequential to a junior executive. Attitude and morale studies uncover many such sore spots which can be easily and quickly remedied, thus improving labor relations within the plant, increasing production, and often forestalling more serious trouble.

INDUSTRIAL ENGINEERING

While the activities of the human and the industrial engineer overlap to a degree, their major interests diverge along two lines. The primary interest of the industrial engineer has been in the machine, or in the arrangement of machines to insure an optimal flow of materials being processed. Interest in the operator exists but has been forced into a secondary position by the character of the engineer's training. Instruction in psychology and in the social sciences at schools of engineering is a relatively new development. There are a few engineering schools in which these disciplines have taken root, but the number is increasing every year. The most serious limitation of the industrial engineer in the field of human engineering, however, is the lack of an experimentally verified literature which could be applied to human problems in industry. The engineers are aware of their limitations in this field; the current mechanical engineering literature contains many searching analyses of these shortcomings. *Mechanical Engineering*, 1946, contains no less than twenty-two articles covering various aspects of the human problem in industry; the engineers' inadequacy in meeting these problems was the principal theme of many of them. In no one of these was there a report of the results of an experiment such as is commonplace in the psychological literature.

The difference in approach, then, between the industrial and the human engineer is one of degree, one of emphasis on men rather than on machines, and on the development of a useful experimental literature.

Naturally, the overlap in approach of both the industrial and human engineer may raise concern among some of them. This concern is groundless. Engineering schools have a difficult enough job of training a good engineer in five college years and psychologists have found seven or eight years all too short a time to train a good psychologist. Ultimately, the engineer and the psychologist will have to work together at least in the bio-technical field. The technical scope is too vast for one type of professional person. It would be more practical and constructive for the psychologist and the engineer to appreciate each other's technical services and understand his own limitations.

PROFESSIONAL TRAINING FOR THE HUMAN ENGINEER

The growing field of human engineering is still too young to warrant the recommendation that a special curriculum be organized leading to a degree in human engineering. The student interested in this field would do wisely to prepare thoroughly in physics and take basic courses in chemistry and human physiology at the undergraduate level. Courses in statistical methods, business and industrial, experimental and physiological psychology, as well as courses in experimental design, human motivation, and bio-mechanics, if available, should receive the highest priority. An internship with a practicing organization would be invaluable.

CONSULTATION VERSUS EXPERIMENTATION

A human engineer can be invited to contribute to a design problem in the capacity of a consultant, as a person with a definite fund of precise information pertinent to a specific problem. Where gaps in knowledge exist, he may be asked to make an educated guess, where such a guess would be better than an uninformed one. For example, knowing that colored objects in uniform illumination of the same color tend to become relatively whitish (appear to lose their saturation), the consultant can immediately recommend the elimination of red markings on dials that would be used in red illumination. He would know that such markings would not stand out well enough against a white background.

More often, however, research is needed for the proper solution of a design problem. Experimental research may be necessary in a specific area of limited applicability. A comparison of the efficacy of two proposed pieces of equipment is an illustration of this type of experiment. On the other hand, the research may be of a basic type, such as the determination of the plane in which crank motion is fastest, or the size of the knob in which the fastest adjustments may be made with accuracy. Such information may be applied to a variety of design problems. Information on the effects of ozone, or of ionized air on human performance, in the concentrations prevailing in the upper atmosphere, is similarly basic information.

The extension of basic research is the lifeblood of the continued development of the field of human engineering. Where will the research come from and who will support it? A substantial amount, it is reasonable to expect, will continue to come from the university psychological laboratories in the form of master's and doctoral theses and from staff members with a human engineering interest. The Army and Navy have embarked on a far-sighted program of supporting basic research not necessarily related to their immediate needs. Such a program ultimately will require the sympathetic understanding of the average American to insure continued support.

Industry stands to gain considerably by the development of bio-mechanics. It should, therefore, share part of the burden of its development. Such moral responsibility can be assumed by an industrial agency in specific contracts by insisting that, to the extent that it is practical and pertinent, as much basic research be done as an operating budget will permit. From a practical point of view the psychological agency performing the human engineering service cannot insist on such a program, it can only point to its social wisdom.

In summary, a need exists for psychologists who can apply their training to the field of bio-mechanics, particularly in its industrial setting. Its growth will depend on the extent to which basic research will be supported by government agencies, universities, and by industry itself.

READING REFERENCES

Chapanis, A.; W. R. Gardner; and C. T. Morgan, *Applied Experimental Psychology*. New York: John Wiley & Sons, Inc., 1949.

Coakley, J. D., *Human Influence on the Product of Automatic Machines*. New York: Bio-Mechanics Division, The Psychological Corporation, 1947.

De Haven, H., "Mechanics of Injury under Force Conditions." *Mechanical Engineering*, 1944, **66**, 264–268.

DeSilva, H. R., *Why We Have Automobile Accidents*. New York: John Wiley & Sons, Inc., 1942.

Dunlap, J. W., "Men and Machines." *Journal of Applied Psychology*, 1947, **31**, 565–579.

Fitts, Paul M., *Psychological Research on Equipment Design*. Washington, D.C.: Government Printing Office, Army Air Force Psychology Program Research Report No. 19, 1947.

———, "Psychology and Aircraft Design; A Study of Factors Pertaining to Safety." *Mechanical Engineering*, 1947, **69**, 135.

Haire, Mason, "Psychological Factors in the Design of Equipment." *Mechanical Engineering*, 1947, **69**, 399.

Hooton, E. A., *A Study in Seating*. Gardner: Heywood-Wakefield Co., 1945.

Kappauf, W. E., "History of Psychological Studies of the Design and Operation of Equipment." *American Psychologist,* 1947, **2**, 83–86.

Lewisohn, S. H., *Human Leadership in Industry; The Challenge of Tomorrow*. New York: Harper & Brothers, 1945.

Luckiesh, Matthew, *Light, Vision and Seeing*. New York: D. Van Nostrand Company, Inc., 1944.

McFarland, Ross A., *Human Factors in Air Transport Design*. New York: McGraw-Hill Book Company, Inc., 1946.

Roethlisberger, F. J., and W. J. Dickson, *Management and the Worker*. Cambridge, Mass.: Harvard University Press, 1939.

Taylor, F. V., "Psychology at the Naval Research Laboratory." *American Psychologist,* 1947, **2**, 87–92.

CHAPTER VII

TRANSPORTATION

50. Street and Highway Traffic	Theodore W. Forbes
51. Railroads	W. Fred Cottrell
52. Aviation	John C. Flanagan

50. STREET AND HIGHWAY TRAFFIC

THEODORE W. FORBES, PH.D.

Chairman and Associate Professor of Psychology, University of Hawaii; Traffic Consultant. Formerly: Research Associate on War Projects at Harvard University, University of Pennsylvania, and Stevens Institute of Technology; Assistant Professor, Bureau for Street Traffic Research, Yale University; Instructor, National Institute for Safety Training; Consultant on Accident Analysis and on Highway Sign Research; Fellow and Research Associate, Harvard Traffic Bureau; Research Psychologist, New York State Psychiatric Institute and Lecturer in Psychology, Barnard College, Columbia University; Instructor in Psychology and Research in Driver Tests, Ohio State University.

The province of psychology in street and highway traffic is the study of the reactions of auto drivers and pedestrians, as well as those of bus and street-car operators in relation to safe and efficient traffic. Moreover, the attitudes of the *general public* toward various traffic facilities and traffic regulations and their understanding of these facilities and regulations must also be included since, for success, the general public must back these measures.

IMPORTANCE OF THE HUMAN FACTOR

In 1924 the First National Conference on Street and Highway Safety focussed attention on the traffic problem on a nation-wide scale. Various investigators pointed out, as the result of accident studies, that a small number of drivers sustained many more accidents than the remainder (Bingham; Baker; Lauer & Forbes in Weiss and Lauer, 1931, & others), and that only from 10 to 25 per cent of the accidents could be attributed to physical causes.

Annual figures on accidents for the United States continue to show that the largest proportion of accidents are caused by human rather than physical factors (National Safety Council, *Accident Facts,* 1937.) In fact, a recent report on one of America's most up-to-date highways indicates that, although accidents have been greatly reduced, the human element is still the major factor in the accidents which occur on this highway (Baldwin, 1946).

RECOGNITION IN TRAFFIC ACTIVITIES

Driver psychology has received most attention in connection with the safety problem, but the importance of the human factor has also been recognized in all the various fields of traffic activity. These fields include traffic regulation and enforcement, design of highways, signs, signals and highway markings, driver licensing, and selection and training of operators of street cars, trucks, and buses. In the field of highway safety education, also, wide use has been made of the results of psychological studies and of psychological techniques of education and advertising.

DRIVER CHARACTERISTICS

ANALYSIS OF ACCIDENT DATA

Numerous studies of accident data have been made in an attempt to determine the characteristics of drivers who have accidents. Only a few of these studies, however, present figures on the relative size and exposure of various driving groups, from which proper correction and evaluation of the results can be made.

Age and Experience. In a study of 29,531 Connecticut drivers, the accident records were analyzed by 5-year age groups and compared with the number of that age in the total group of registered drivers. The youngest drivers (20 and under) sustained 50 per cent *more* than their share of accidents and the rate fell steadily until the 35- to 40-year group showed 20 per cent *less* than their share. This rate was maintained up to sixty years, beyond which there was a slight increase. These facts were corroborated in a study of a group of about 11,000 Washington, D.C., drivers (Johnson and Cobb, 1938). The tendency was even greater for drivers in fatal accidents.

DeSilva (1942) also reports a similar high accident rate among young drivers, and reports that mileage was greatest for 25- to 29-year men and less for younger and older groups. Thus, the high ratio of accidents among younger drivers cannot be due to greater mileage and exposure.

A higher accident rate has been found also among inexperienced commercial operators. (Kraft and Forbes, 1940, unpublished.) In this study some of the inexperienced operators were found to be older men, which suggests that experience may be the more important factor.

APPRECIATION OF HAZARDS. A careful case study of 50 high accident motormen (Cleveland Railway Co., etc., 1929) resulted in the conclusion that "failure to recognize potential hazards" was the second highest primary cause of accidents (12 per cent of the cases). Here again men with under five years of service showed a higher accident rate than those with from five to ten years of experience. Although this relationship might be an artifact in a group of transit operators, since high accident men are more likely to be separated from the job, it should not be so among private drivers. Incidental unpublished observations (Lauer and Forbes, 1927–29; DeSilva and Forbes, 1937; Kraft and Forbes, 1944) indicated that the experienced driver or bus operator learns to anticipate hazards from minor cues which may not be observed or appreciated by the novice. For example, the experienced driver learns to observe the front-wheel attitude of parked cars along the side of the street as a warning of impending pull-out into traffic.

In similar fashion, the increase of stopping distance with higher speeds may not be appreciated by the novice. Significantly, Johnson (1938) also found 2.3 times the expected number of from 16- to 25-year-old drivers in a group of 1,135 arrests for speeding in the District of Columbia.

ATTITUDE OF THE DRIVER. Although few large scale studies of driver attitudes have been published, their importance should not be overlooked. The case study referred to above (Cleveland Railway Co., etc.) placed "faulty attitude" (14 per cent), "impulsiveness" (10 per cent), and "irresponsibility" (8 per cent) among the five most frequent primary causes. Other case illustrations of the apparent relationship of faulty attitude to accidents have been given (DeSilva and Forbes, 1938). Indirect evidence of the importance of attitude factors is also available. Analysis of law violations by drivers in accidents showed "violation of right of way" as the largest single violation when the average of all accidents in twenty-seven states in 1944 was obtained (National Safety Council, *Accident Facts,* 1945).

ATTENTION. A study by Ringwald (Weiss and Lauer, 1931) suggested that the better drivers did not center attention continuously in any one direction, but rather alternate it rather evenly in all quadrants.

MOTOR-COORDINATION AND SKILL

"DRIVER TESTS." Attempts to measure driving skills started in the United States with the use of laboratory tests by Munsterberg (1912), followed with other tests by Wechsler (1925) and Snow (1926) of taxi drivers. Moss (1925) measured reaction times on the road. Viteles (1925) used tests requiring complex coordinations with motormen and Shellow (1926) and McCarter (1931) extended this work.

Lauer (1928–31) used tests of complex eye-hand-foot coordination, employing automobile controls, a "steering" task involving control of a pointer moving laterally, and also reaction to signals. Forbes (1928–31) developed a Miniature Highway test involving motor coordination and the estimation of the relative velocities of opposing streams of traffic presented in miniature on a small-scale road (40′ by 6″ endless belt). The visual effect was similar to that presented to the driver as he moves along the road (Weiss and Lauer, 1931).

The Miniature Highway test was later modified and improved in collaboration with DeSilva and included in the Harvard Traffic Bureau test battery (DeSilva, 1935; DeSilva and Forbes, 1937).

Lauer later modified his coordination test by introducing a "roadway" depicted on a revolving drum and adding a general ability, a glare vision, and other visual tests to form the Iowa State College battery (Lauer, 1936).

DeSilva developed tests using automobile pedals and steering wheel and a scene depicting a roadway which moved laterally in front of the subject. The subject was required to follow this movement by use of the steering wheel. The Brake-Reaction-Time test required the subject to depress the brake pedal in response to a red signal. The vigilance test required both steering coordination and response to occasional red signals simultaneously (DeSilva, 1935). A speed estimation test was also developed using miniature cars moving at constant speeds across the subject's line of vision, left to right, and in addition a test for recovery from glare. This test measured acuity during glare and also the time for recovery after exposure to glare. Peripheral vision and color vision tests and the Miniature Highway test were added to form the Harvard Traffic Bureau battery (DeSilva and Forbes, 1937).

A test for motormen involving motor coordination and reactions to complex signals was developed by the Pittsburg Street Railway Company (1935). The subject was required to use different combinations of steering-wheel movement, clutch-and-brake pedal reactions to different combinations of light signals. This test has been improved and standardized for use by transit companies (Cleeton, Kraft and Royster, 1946).

Paper-and-pencil tests of attitudes and knowledge of the car and of traffic regulations have been used among private drivers (Lauer, 1936; Abercrombie, 1941) and commercial bus operators (Cleeton, Kraft and Royster, 1946). Group tests of driver information were developed during the war for army truck drivers (Staff, AGO, 1943).

A standardized road test has been used as a criterion of driving skill for private drivers (Lauer, 1936), for licensing drivers (DeSilva, et al., 1939), and for army truck drivers (Staff, AGO, 1943).

RESULTS FROM DRIVER TEST STUDIES

MAGNITUDE OF REACTION TIMES. Simple-brake-reaction time (i.e., the time required to depress the brake after a signal flashes when the subject is attending to no other task) was shown to average from 0.4 to 0.5 second. However, when the driver was required to carry out a steering task simultaneously, the reaction-time average increased to from 0.6 to 0.7 second (DeSilva and Forbes, 1937). In actual driving on the road, the performance required of the driver is even more complex and the time required for a driver's reaction may be still greater.

REACTION-SPEED, SKILL, AND DRIVER'S AGE. A report from New England showed that drivers begin to show slower reaction somewhere between the age of thirty and forty years on the average. Ability to see against glare showed a similar decrease as did certain test scores involving both skill and speed of reaction. This decrement was more marked among groups of accident-repeaters than among volunteer groups of drivers (DeSilva and Forbes, 1937; Forbes, 1938). There was, of course, considerable variation from the average among *individuals* in each age group, i.e., many older drivers made better scores than certain younger ones.

TEST SCORES AND ACCIDENT EXPERIENCE. Two general methods have been used for evaluating the relationship of test performance and accident records. The first of these is comparison of the average score of a group of accident-repeaters with that of a group of nonrepeaters, evaluating the difference of the means for statistical reliability by well known statistical methods. Unless statistical reliability is proved the difference of group averages may be merely due to chance.

The second method uses a single group of drivers chosen to represent a cross-section of all drivers. It evaluates the relationship between test scores and accident rate for the various individuals by means of a statistical coefficient of correlation. In both methods, there is needed, not only proper statistical procedure, but also the use of accurate reasoning. As pointed out in a recent critical analysis, many of the reports on this subject have not been based on valid evaluation techniques (Johnson, 1946).

Several investigations using the group comparison method have shown evidence that many of the tests measure factors which are *to some degree* related to accident experience. Lauer (1936) reported that poor scores in (*a*) manipulation of controls, (*b*) ability to observe signs, and (*c*) time required to complete the "trip" on the Iowa State coordination apparatus showed a statistically reliable relationship to accidents. From the Harvard Traffic Bureau battery, reaction-time and scores on the steering and vigilance-steering tests and combinations of these showed statistically reliable differences between groups (Forbes, 1938; DeSilva and Forbes, 1937).

A number of other scores from the latter study show group differences which are not quite reliable statistically (a probability of 2 in 100, or less of being due to chance). Logically this might mean that the factors measured had no relationship to accident. On the other hand, if they had such a relationship in certain individual cases, but not in the majority, such a result might be obtained. Evidence from case studies seems to indicate the latter alternative. In this category were scores on glare vision, visual efficiency, and miniature highway "safe passes."

By the correlational method, an extensive evaluation of the Iowa State and Harvard test batteries was conducted by the Highway Research Board in 1937. Accident records for the preceding six years were obtained from Connecticut records, which had been chosen as representing the best available. Both test scores and accident records were obtained on 3,161 drivers.

It was found that the 18 per cent of the drivers who earned the *highest* scores on the group of tests had less than *half* the accident rate of the others, and the 17 per cent who earned the *lowest* scores had an accident rate *three times* greater than the others (Johnson, 1946). Thus, the existence of a relationship between the abilities tested and accident experience was corroboratd, and the usefulness of the tests for *prediction on a group basis* was indicated.

Correlations between scores of any individual test and accident rates, however, were consistently low and the highest correlation coefficient obtained between weighted multiple scores and accident rate was about 0.35. This means that prediction of accidents for *individual operators,* on the basis of these tests scores only, would not be of practical accuracy (Johnson, 1946).

MEDICAL TEST DATA AND ACCIDENT RATE

Medical data also show rather low correlations with accident records. Certain physical data, such as visual acuity, color vision, and blood pressures, were included in the low correlations reported by Johnson above.

Another study of 482 street-car operators showed similar low correlations between accident rates and medical data. The accident records maintained by the company were more complete than the public records, and accident rates were corrected for varying exposure to hazards among the operators (Kraft and Forbes, 1944). It is of interest that certain measurements, of which blood pressure was an example, appeared to exhibit a *curvilinear* rather than a *linear* relation to the accident index, i.e., very low as well as high blood pressures tended to be associated with higher accident rates. However, the correlation ratio (measure of nonlinear relationship) was too small to be of practical use for prediction.

This low relationship in two studies of large groups is of importance in evaluating other studies of small groups which have reported a startlingly high relationship (e.g. Brodie, 1941, 26 drivers in each group). Apparently, in selecting the groups an unintentional factor must have entered which resulted in the preponderance of low blood pressures among the accident repeaters of this study. In fact, the same group of investigators has reported a study of a larger group of drivers showing results similar to those of Kraft and Forbes (Center for Safety Education, 1948).

ROAD TESTS AS MEASURES OF DRIVING SKILL

Several investigations have employed a standardized road test in which observers check specific items of good and poor operation by the driver (Lauer, 1936; Abercrombie, 1941; Staff, AGO, 1943), in order to obtain ratings of actual driving skill.

In one study (Staff, AGO, 1943) examiners not only checked 47 specific items of operation but also gave an *over-all* rating to each driver. Data was analyzed from 1717 road checks made by specially trained examiners. It was found that the total of "unsatisfactory" checks correlated as high with the over-all rating ($r = 0.83$) as did the best weighted score from selected items. Furthermore, although the split-half reliability of the items' check was high (0.83), the reliability based on items checked by *different examiners* was considerably lower ($r = 0.53$). However, the reliability of *over-all ratings* of the same men by different examiners was fairly good ($r = 0.72$).

The use of a check list consisting of specified items, for example, "uses wrong gear uphill" is, of course, desirable in order to standardize the basis of judgment but apparently the particular errors made were of less importance for the purpose of obtaining consistent ratings than the total picture.

OTHER EVIDENCE ON DRIVER ABILITIES

ESTIMATION OF SPEED. There is fairly good evidence that judgment of speed and distance on the road by most drivers is not very accurate or reliable. In a study in which groups of subjects rode in one car and judged the location at which an oncoming car would pass, retest reliability for two series of judgments was only 0.14 ± 0.13. Incidental observations suggested that the subject's expectation of the speed of the test car sometimes affected his judgment more than visual impressions from the approaching car (Forbes, 1932).

Another study corroborated the inaccuracy of judgments of speed and distance (Forbes and Matson, 1939). A total of 785 passes on the highway were photographed and timed. Analysis showed that 10 per cent of the drivers observed misjudged the speed of the approaching car sufficiently to force one of the other drivers over onto the shoulder to avoid collision. A further evidence of the inaccuracy of these judgments was obtained in the same study by analyzing cases where a refusal to pass was followed by a pass. The opportunity refused was more than sufficient about as frequently as it was less than sufficient for the pass made subsequently by the same driver.

ACCIDENT PRONENESS

Many investigators of the accident problem have pointed out that a small proportion of the drivers have a large proportion of the accidents. About

7.5 per cent of drivers had 50 per cent of the accidents (Weiss and Lauer, 1931); 3.9 per cent of the drivers had about 36 per cent of the accidents (Secretary of Agriculture, 1938). The term *accident prone* has been quite generally applied to this group of drivers.

It has been suggested that if these drivers could be either ruled off the road or made into better drivers, the most effective accident reduction would result. This point of view is logical provided that these drivers are *consistently* more liable to accident than others. It is clear, however, that *some* high accident drivers are "just unlucky."

We thus have the problem of distinguishing the truly accident-prone drivers from those to whom chance brought a high number of accidents. One method proposed for determining the number of such high-accident-by-chance drivers has been the use of the statistical distribution derived from the Poisson series (Secretary of Agriculture, 1937; Johnson, 1946). The Poisson formula, like any such function, however, is based on certain definite assumptions as to the hazards to which the various members of the group are exposed.

Another method of defining consistently accident-prone drivers has been suggested, namely, that they should be those having two or more accidents in the first 3 years, and one or more accidents during the second 3 years over a 6-year period. On this basis it was shown that less than 1.3 per cent of the 29,531 Connecticut drivers would qualify (Forbes, 1939). This percentage is *smaller* than the expected figure obtained by the Poisson formula. Also this group had only 3.7 per cent of the accidents in the second 3-year period.

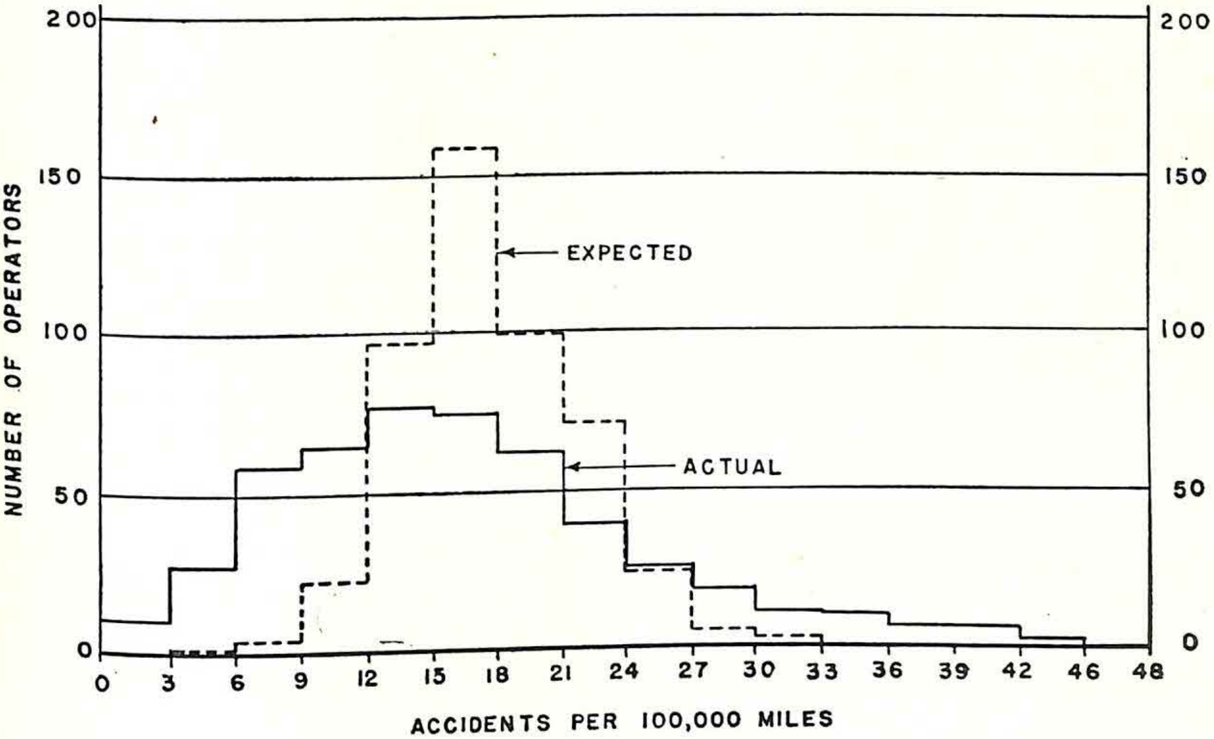


Figure 1 (VII–50): *Comparison of Mileage Rates and Transit Traffic Accidents; Expected Rates vs. Actual Rate. (For 482 Street Car Operators during Three Years, 1937–1939.)*

Still another approach has been tried. In a study of street-car operators, accident experience was available for different runs among which men had been rotated. The over-all accident hazard for different hours of the day was also available, as was an estimated mileage represented by each man's trips. On the basis of these and the assignments of men to hours of work and to runs, an expected accident figure was computed for each man. The distribution of these individual accident expectancies thus gave an *empirical* expected-accident distribution, which turned out to be a symmetrical one which was quite different from the Poisson curve (Kraft and Forbes, 1944). This is shown in Figure 1.

The best method of defining and estimating the relative importance of the truly accident-prone driver is thus not yet clear. It is clear, however, that many high-accident drivers are not consistently accident prone over long periods and there is considerable evidence that many of them can drive safely after having a series of accidents. It is true, however, that on a group basis there is a higher probability of ensuing accidents among high- than among low-accident drivers, as Table 1 shows.

Table 1: Accident Rates in Following Three Years of Driver's Group According to Accidents in Preceding Three Years

Accidents per Driver	
1931-33	1934-36
0	0.10
1	0.20
2	0.30
3	0.48
4	0.70

From a study of 29,531 Connecticut drivers (Secretary of Agriculture, 1938).

THEORY OF ACCIDENT CAUSATION

It can be concluded from the foregoing and similar studies that accidents chargeable to psychological factors do not result from any one or two characteristics of the drivers. On the contrary, exposure to hazards varies widely and the occurrence of a critical situation, and errors of skill, judgment, attitude, and similar factors must be simultaneous to cause an accident. Thus a person may drive safely in spite of limited skill by avoiding hazardous traffic conditions. Furthermore, errors on the part of *two or more drivers simultaneously* may be necessary to cause an accident. Observations on the road have shown numerous instances where foresight and skill on the part of two drivers saved from accident a third driver who made a definite error of judgment (Forbes and Matson, 1939).

Such a conception of accident causation is in accord with the general findings of a consistent but low relationship between accident rate and most of the driver characteristics measured. It is also in accord with the findings (Johnson, 1946) that factors, such as restricted peripheral vision and poor glare vision, showed no apparent statistical relationship with collisions from the side and night accidents, although at the same time "clinical" case studies indicated that in certain cases such deficiencies were probably causal.

SELECTION OF OPERATOR

It has been estimated by Viteles (1932) that accidents cost the taxicab industry 22 million dollars in one year, and the Boston Elevated Railway Company charged over one million dollars to injuries and damages during 1928–1929. Over and above this is the cost of training new operators when others prove unsatisfactory.

Efforts have therefore been made to select new operators, both by means of tests and of interviews. Recently a series of tests and a standardized interview have been developed for selection of street-car and bus operators by the American Transit Association (Cleeton, Kraft, and Royster, 1946). Checks of validity indicated that those passing above a certain recommended score produced 100 successful out of 150 total trainees on one test, and 100 out of 130 on a second. Those testing below a recommended unsatisfactory score produced only 100 successful out of 280 on the first and 100 out of 170 on the second. A Personal Reaction Test is recommended for clinical use by interviewers, but not on the basis of a pass-or-fail score.

During the war a group test of driver information was developed for use in selecting Army truck drivers (Staff, AGO, 1943).

DRIVER AND PEDESTRIAN EDUCATION

To obtain rapid and safe highway transportation the general public needs to be taught proper driving and safety practices as well as safe walking habits. This knowledge can be presented through various mediums including advertising posters and brochures, exhibits at auto shows, radio skits, programs put on by local auto and safety organizations, driver training for adolescents and adults, and retraining "clinics" under the auspices of police and industrial fleets.

GENERAL SAFETY CAMPAIGNS

A National Traffic Safety Contest (National Safety Council, 1945) has been carried on for a number of years through the cooperation of national and local safety council organizations and state and local motor vehicle and police officials. This contest involves the keeping of standardized accident records by states and cities, and monthly reports of the figures. Awards are given for the best record of improvement shown by states and by cities in each of several population classes. Similarly, fleet safety contests are carried on by local safety councils through the cooperation of local industrial and business concerns operating fleets of vehicles.

EDUCATIONAL USE OF DRIVER TESTS

A battery of driver tests (Harvard Traffic Bureau battery) was operated in 1936 through 1938 as an exhibit for the public at the New York Automobile Show and the public showed great interest. Beginning at about the same time various automobile clubs, insurance companies, and similar organizations started the use of similar tests as educational devices.

Unfortunately, in their later use, unstandardized tests and advice given by nonpsychologically trained personnel sometimes appeared to show a lack of appreciation of the limitations of the tests. If properly standardized and handled, however, the tests can be of value in calling attention to individual differences in abilities and skills and the significance of psychological factors for safety.

DRIVER TRAINING

The teaching of effective and safe driving habits, together with information about the automobile and a knowledge of traffic regulations, has been introduced into the high schools in various states and cities. The value of such courses is indicated by reports from Cleveland, Ohio, and State College, Pennsylvania, that students who had had both classroom and road training were involved in only *half* as many accidents as comparable groups of students who did not have this instruction. From Delaware it was reported that the untrained groups showed 46 per cent more violations (including accidents) than the trained drivers (National Commission on Safety Education, 1947).

A standardized technique for training was developed by Neyhart and a training course for teachers was started at Pennsylvania State College in 1936 (DeSilva, 1942). Such teacher training courses have been sponsored since then by the American Automobile Association and other safety organizations, and administrative manuals and textual materials are available (National Commission on Safety Education, 1947; American Automobile Association, 1937).

A report on the results of training indicates a substantial gain from instruction both in tests of knowledge and of skill, even on the part of those who were already drivers (Abercrombie, 1941).

REEDUCATION OF DRIVERS

THE CLINICAL METHOD IN ACCIDENT PREVENTION. In a number of instances reeducation of vehicle operators in industry has been accomplished by use of the clinical or case-study method. Viteles (1932) presents the case well for the clinical method, pointing out that it allows the investigator to examine the whole individual rather than isolated characteristics. Viteles summarizes reports by Bingham, Slocombe, Viteles, and others in which economic savings of as much as $300,000 in accident costs in one year are indicated.

DRIVER TEST CLINICS. A battery of tests was used as a part of a clinical procedure on groups of repeaters obtained through the cooperation of police departments in various Eastern states in 1936–37 by the Harvard Traffic Bureau. An interview followed the test, and often the driver, himself, volunteered suggestions as to how a characteristic in which he had scored low might have played a part in a certain one of his accidents. In other cases, such a relationship appeared probable in certain accidents when reviewing the case history (DeSilva and Forbes, 1937).

Among two groups follow-up research on accident records showed a reduction of accidents which was considerably greater for the clinic group than for accident-repeaters not attending the clinic. In one comparison, accidents in the succeeding nine months were 13 per cent of those in the preceding period for clinic repeaters as against 32 per cent for repeaters who did not attend (Forbes, 1938). The well known tendency for a high-accident group to appear to improve due to chance factors ("regression toward the mean") is seen in the second group but there was a further gain of 19 per cent among the clinic group.

Johnson (1946) estimated the amount of improvement reported by Slocombe and Brakeman due to "regression toward the mean." He found that there remained over and above this a gain of about 22 per cent due to reeducation, which is of the general order of that noted above.

Driver test clinics have also been used for the reeducation of law violators as part of a violator's school.

From the practical point of view, such clinical use of driver tests is expensive partly because of cost of equipment, but even more because of the cost of qualified personnel if continued for a long period.

Whether such use is justified would seem to depend on the particular situation.

PEDESTRIANS

A summary of pedestrian studies (American Automobile Association, 1939), points out that the largest number of pedestrian deaths occur among children under fifteen years of age and among persons over fifty years of age. Programs for the protection and education of school children through Junior Traffic Police and similar activities have been promoted, and the traffic deaths among children have shown a downward trend. Evidence is presented that in terms of the number of persons of a given age-group in the population, persons over fifty were 5 or 6 times more likely than school children to be killed as pedestrians.

Several causal factors are suggested as the result of studies from several states. One of these is unfamiliarity of older pedestrians with motor-car operation, and the difficulty of stopping and of seeing pedestrians at night. Another is drinking on the part of pedestrians. Unsafe walking habits also probably play a major part. This report represents one of several efforts to educate the public with regard to pedestrian safety.

REGULATION AND ENFORCEMENT

LICENSING TESTS

Motor vehicle departments of many states administer tests to drivers before giving them a license to drive on the public highways. The examinations have been varied but all have included a practical road driving test and a rough test of vision.

DeSilva and others (American Association of Motor Vehicle Administrators, 1939), have suggested standards for state driver tests and the methods of administering them. The procedure recommended includes instructions for visual tests, specifications for a standard road test, and short answer written examinations of traffic knowledge and of traffic laws.

FATIGUE AND REGULATIONS

A number of studies have been made of fatigue from riding in and driving automobiles. Baker (1935) investigated driving conditions in relation to fatigue of commercial truck drivers on long routes. The need for shorter runs and for better sleeping conditions in sleeper cabs was indicated. Ryan and Warner (1936) made various psychological and physiological measurements of bus drivers before and after their daily 8-hour trip behind the wheel, and found a marked increase in errors in certain psychological tasks. (Lawshe, 1939.)

On the basis of available knowledge and expert opinion, the Interstate Commerce Commission, in its regulations for commercial bus and truck operators issued in 1939, reduced markedly the number of hours at a stretch that a driver could be required to work. (Lawshe, 1939.)

Definite effects of fatigue from truck-driving have been shown by a battery of psychological and medical tests given to 889 truck drivers in a study by the US Public Health Service (Jones, Flinn, Hammond, et al., 1941).

ALCOHOL

Driving under the influence of alcohol is essentially a psychological problem since alcohol affects the accuracy of judgments and skilled manipulations such as those involved in driving (Poffenberger, 1941; also unpublished observations, Harvard Traffic Bureau). The seriousness of the problem has been widely recognized, but strict enforcement has been difficult because one psychological phenomenon was often offset by another. The earlier and rougher measures for intoxication consisted of such tests as being able to walk a line, or the ability to say difficult word combinations. Unfortunately, the human system is capable of mobilizing its forces so that upon arrest a man who previously would not have passed these tests might be able to perform them fairly creditably. Chemical tests of blood, urine, and breath have been utilized to overcome this difficulty. These tests have been applied to gain convictions in a number of states and have been upheld. The highest validity has been found for the percentage of alcohol in the blood. A concentration of 0.15 of one per cent of alcohol in the blood has been recommended as a practical indicator of intoxication (Committee on Tests for Intoxication, 1940).

SELECTION AND TRAINING OF TRAFFIC OFFICERS

Since about 1936 specialized training has been carried on at the Northwestern University Traffic Institute for selected police officers to fit them for responsible traffic police positions in their state or local organization. This training has been sponsored by the International Association of Chiefs of Police and backed by various other organizations and foundations.

Fellowships are offered for which men are carefully selected on the basis of ability and personality by means of tests and interviews. Futhermore, attention is given in the course of training to important psychological factors, such as public relations, personnel, accident analysis, tests for alcohol, safety education, and practical psychology.

MOTORIST REACTION

Stress has been laid on the need for favorable motorist reactions toward traffic regulations in order to obtain the best traffic enforcement. If the motorist feels that a law or a regulation is a reasonable one, proper enforcement of it is relatively simple. For this reason, traffic enforcement programs now include education of the public as to the reasons for traffic laws and regulations. All of the methods used in educational and safety campaigns are employed and at the same time an attempt is made to see that the laws and the regulations actually are reasonable. The individual traffic officer is trained to point out to the motorist the reasonableness of the regulation he has violated.

TRAFFIC ENGINEERING

The traffic engineer might be called the operating engineer of the highway or street system. It is his responsibility to see that all engineering features of the roadway or the street contribute to safe, efficient, and satisfactory movement of all types of traffic. A wide range of requests and complaints on traffic matters comes to his office, so that his work requires the solution of problems involving both physical and human factors, and is thus a combination of engineering and psychology.

ACCIDENT ANALYSIS

Traffic engineering takes much of its direction from the analysis of accident records. Such analyses are carried on by such agencies as the National Safety Council, by state and city traffic engineering departments, and by police departments. These analyses involve (1) physical factors, such as roadway condition and design, and the visibility of signs and of signals; and (2) psychological factors, such as recklessness, speed of operation, driver's condition, and the reaction of drivers to factors of highway design.

As noted earlier (see Accident Proneness), there is a small group of individuals who sustain an unusually high number of accidents, but the consistently accident-prone group are responsible for only a relatively small proportion of the total accidents on the highway. Accident analysis for the traffic engineer, therefore, largely concerns factors influencing all drivers.

Driver attitudes may play an important role in accident records of particular highways, as is illustrated by the following instance: two military truck convoys, roughly identical in makeup, started from the same point for summer manoeuvers of a National Guard unit. One traveled a very rough route full of blind turns, hills, and narrow roads and the second traveled about the same distance over flat, straight, concrete roads. The latter had a number of accidents and the former none, apparently due to the alertness of its drivers to the hazards of the route.

DRIVER TIME-FACTORS

Driver time-factors may be considered as reaction time, perception-reaction time, and judgment-perception-reaction time, for convenience, designated as reaction time, perception time, and judgment time. Leading highway design engineers have become conscious of the importance of these factors in highway design.

The reaction time interval is illustrated by the time the driver takes to react to a red traffic signal or an obvious danger which calls for an emergency stop. Perception-reaction time involves reactions to more complex situations, while judgment time involves both perception and judgments in complex situations. Values of 0.50 and 0.75 second may be used as practical average times for the first two, whereas for judging speeds of oncoming cars in traffic average values of 1.0 to 1.5 seconds are involved. The exact time-value will depend upon the complexity of the traffic situation, and, in fact, the time-interval required may be as long as 2.8 to 3.2 seconds (Forbes, 1939).

Knowing these time-values and the speeds for which he is designing, the engineer can provide sufficient warning distance or clear sight-distance at critical points. At complex intersections he should also keep the driver's task as simple as possible, since, not only the above time-values, but also a period of 0.6 to 0.8 second (or a multiple of this) for moving the eyes and fixating a new object must be provided (Forbes, 1940).

VISIBILITY

Seeing at night by street lights raises numerous problems of visibility which have been studied with respect to design of lighting equipment. Under these conditions a large part of seeing is by silhouette and, therefore, uniform *pavement brightness* rather than uniform *illumination* is necessary. Glare from overhead illumination may reduce visibility by as much as 50 per cent (Reid and Channon, 1936).

HEADLIGHTING

Seeing at night when driving with headlights alone involves other problems of visibility, such as the color of highway surface, of objects, and of pedestrian clothing. It has been shown that materials having low reflection reduce visibility, and also that glare from on-coming headlights may reduce perception distance by as much as 70 per cent. Further, the motorist may see an unexpected obstacle (dummy pedestrian) at only about 50 per cent of the distance at which he sees an expected one. With different headlight intensities visibility distances were also measured and the results indicated that modern speeds could not be used with safety with older headlights, but that, with the modern brighter headlights, it was just possible to stop safely after seeing an unexpected obstacle (Roper and Howard, 1937).

HIGHWAY SIGHT DISTANCES

The distance ahead that is clear for vision at any point on the highway is called sight distance. Sight distance is a major consideration in the design of modern highways. It must always be adequate for emergency stops at the speed for which the highway is designed, and on two-lane highways there must be at intervals sufficient sight distance for safe passing.

Observations by Matson and Forbes (1938) of the passing behavior of drivers on highways confirmed the opinion of many traffic and highway engineers that passing distances are generally too short. Not only did the passing distance required show a linear relationship to the speed of the vehicle passed, but these distances were also affected by the driver's attitude, i.e., much shorter passing opportunities were utilized in hilly terrain where the driver did not expect better opportunities. Large-scale studies by the Public Roads Administration reported in the Proceedings of the Highway Research Board have generally corroborated this evidence.

Drivers' judgments of speed and distance of oncoming cars in passing is none too reliable (Forbes and Matson, 1938). In these observations, 10 per cent of the drivers misjudged and forced the oncoming driver to "make three lanes out of two," while others refused to pass when there would have been time and distance enough to do so. Therefore a margin of safety should be provided.

SIGNAL COLORS

Traffic signal lenses have been designed to distribute intensity in the beam toward the driver so as to furnish the best visibility possible, both at a distance and also close up. Considerable attention has also been given to the color of the lenses.

Since it is known that about 8 per cent of the males in the United States are red-green color blind, signal color is of importance. A red and a green containing some yellow and blue respectively have been adopted and uniform placement of the red in the topmost position has been specified (Manual for Uniform Traffic Control Devices, National Conference of Street and Highway Safety). Both the selection of colors and the standard vertical placement of colors aid the color-blind driver in identifying the signal.

Recent standards for signal colors (Joint Committee of the Institute of Traffic Engineers and of Illuminating Engineering Society, 1940) specify colors in terms of the hypothetical average normal observer defined by the International Commission on Illumination.

SIGN EFFECTIVENESS

The "legibility distance" of different standard alphabets for highway signs (the distance at which the letters can be read) was studied for the Pennsylvania Turnpike Commission by Forbes and Holmes (1939). Results were expressed in terms of the height and width of the letter and the visibility distance at which 80 per cent of a large group of observers could read the sign. Curves of these relationships were published. For speeds of 50 miles per hour, a wide letter 10 inches high or a narrow one 12 inches high were recommended as a minimum, unless advance placement of the sign is used.

As a result of the study, letters of from 6 to 18 inches in height were used to fit different speeds and situations on the turnpike. A more recent application of these results to round letter alphabets has been reported by Loutzenheiser (1943).

Legibility distances of designs using arrows was measured in similar fashion. Arrows with a 45-degree angle forming the head, and a shaft length at least as great as the barb-to-barb width proved most effective. (Connecticut State Highway Department and Yale Bureau of Highway Traffic, Forbes, 1941.)

Studies of other factors in signs (Forbes, 1939) demonstrated that the well known factors affecting legibility and attention-value of advertising signs apply, with slight modifications, to highway signs. Thus relative size and color-contrast are of importance for attention-value and, together with position of a sign in relation to others, these determine which sign will be read first. The greatest contrast between letters and background or between sign and background gives the best legibility and visibility. Although studies by the Bureau of Standards in 1933 (Lawshe, 1939) led to the adoption of black and white combinations and black and yellow combinations for highway signs, for attention value the sign selected should contrast with the terrain or the sky against which it is viewed.

SIGN PLACEMENT

It has been shown that, on the basis of psychological data now available, it is possible to determine mathematically the proper size of letter and the proper sign location to use if the driver is to have sufficient time for safe and comfortable deceleration at intersections. The treatment was based on legibility distance data (Forbes & Holmes, 1939), times for reading signs (Lauer, 1932), times for comfortable deceleration (Beakey, 1938), time and number of words read at a glance (Forbes, 1939), the various reaction times discussed in a previous paragraph, and vehicle speed for which the design is intended. Formulae were developed for Case 1 where some cars proceed, others decelerate, but none must stop, and Case 2 where all vehicles must stop (Mitchell and Forbes, 1942).

SURVEY METHODS

For some years, information on the origin and destination of vehicles in city traffic and similar data has been obtained by short questionnaires handed or sent to motorists. The information so obtained has been of importance for the intelligent planning of traffic facilities and arteries. Recently the use of a standardized interview has been adopted by the Public Roads Administration. A sample of dwellings is selected on a geographical basis and each family in the dwelling unit is visited by a home interviewer. It is reported that more valid results are thus obtained (Lynch, 1945).

PSYCHOLOGICAL RESEARCH

Because of the relative newness of traffic engineering as a profession, many psychological factors have not yet received the thorough study that they deserve. Many of them have been handled perforce by the engineers who were confronted with a practical problem and who had to produce a solution with the means at hand. Application of psychological research methods to additional problems in this field should yield further valuable information.

CONCLUSION

The attempt has been made to indicate phases in which psychological studies have contributed toward the solution of highway traffic problems. It has been impossible to treat all of the large number of studies and reports in this field, many others of which deserve mention. It is felt, however, that a fairly adequate coverage, especially of the more recent and better controlled studies, has been given and the reader who is technically interested can obtain from the reviews listed in the bibliography other references which may suit his interest.

READING REFERENCES

American Automobile Association, *Pedestrian Protection.* Washington, D.C.: The Association, 1939.

Center for Safety Education, New York University, *Personal Characteristics of Traffic-Accident Repeaters.* Saugatuck, Conn.: Eno Foundation for Highway Traffic Control, 1948, p. 64.

DeSilva, H. R., *Why We Have Automobile Accidents.* New York: John Wiley & Sons, Inc., 1942.

Lawshe, C. H., *A Review of the Literature Related to the Various Psychological Aspects of Highway Safety.* Lafayette: Purdue University, 1939.

National Commission on Safety Education of the National Education Association, *Let's Teach Driving; An Administrative Guide Book.* Washington, D.C.: The Association, 1947.

National Safety Council, Inc., *Accident Facts.* Chicago: The Council, published annually.

51. RAILROADS

W. FRED COTTRELL, PH.D.

Professor of Sociology and Government, Miami University; President, The Ohio Valley Sociological Society.

Although the railroads constitute one of the oldest branches of industry in America, applied psychology has entered this field slowly. In fact it was only after its usefulness had been established by the street railways that other railroads took it up. Of all the psychological problems that confronted the industry the oldest and one of greatest significance is that of safety.

SAFETY

AMERICAN STREET RAILWAYS

Munsterberg's classic experiment (1912) in the selection of street-car motormen was the first attempt to utilize technical psychology on American railroads. The American Association for Labor Legislation, which had hitherto placed much emphasis upon legislation requiring the improvement of equipment, departed from that emphasis in 1912 and invited Munsterberg to investigate the human factors in railroad accidents. His analysis led to the conclusion that the fundamental psychological problem faced by the motorman was that of responding to multi-plex stimuli. He therefore designed a preselection test capable of discovering men capable of this type of response. His selection test consisted of a glass-covered wooden box, fitted with a cylinder at either end, over which travels a black velvet belt. An aperture in the belt uncovers serially a number of cards representing a street. Street-car tracks are represented on these cards by heavy lines. Potential obstacles to which the motorman must attend are represented by numbers and letters indicating movements of greater or less concern to the operator. The potential motorman moves the belt by turning a crank and calls out the symbols which have significance as rapidly as he is able to discover them. The score on the test is determined by the amount of time taken and the errors made. Munsterberg died before he could make extensive use of this test and McCant's efforts (1922) to supplement it by interview and intelligence test items met with indifferent success.

Munsterberg had neglected to take into account the significance of the coordinated muscular response which must be made in the job situation to the stimuli presented in the test. Viteles (1920) tested both attention and response in a study of motormen for the Milwaukee Electric Railway and Light Company. Viteles' apparatus included signal boards for visual and auditory stimuli, a distraction signal, and an operating stand which supported equipment similar to that used on a street-car by a motorman. When definite standards of learning have been met the potential motorman is given a test in which he responds to signals by the same muscular response that would be required in the actual operation of a car. After modification, permitting automatic recording of reaction time, the apparatus proved to be satisfactory, as shown by Shellow (1924, 1925) and Viteles (1925), who demonstrated that it reduced the accident rate and eliminated many of the men who would otherwise have been likely to be discharged. The rate was reduced still further by therapeutic interview with those men found to have a large number of accidents. These interviews also demonstrated that many accidents are due to temporary emotional difficulties.

Bingham and Slocombe (1927) investigated accidents on the Boston Elevated Railway from the point of view of individual susceptibility to accident. They showed that a large proportion of accidents involve relatively few men, and developed the concept of the accident-prone man. Such men differ from their fellows chiefly in shorter length of service and emotional rather than physical factors. In a few cases physical defects, such as high blood pressure, contributed to accident proneness, but the bulk of the factors found to be associated with the condition were psychological in character. A therapeutic counseling technique was used for the reeducation of accident-prone men, which resulted both in a marked decrease in the accident rate and a sharp increase in operating efficiency. The findings of the Metropolitan Life Insurance Company in studying the Cleveland Street Railway Company (1930) have paralleled those of Bingham and Slocombe and the earlier work of Shellow and Viteles in indicating the presence of accident-prone operators on street railways and the psychological therapy necessary to their rehabilitation. Cleeton (1935) and others have indicated that attention-reaction-time adjustment and general intelligence must be considered in the personnel problems of safety in street-car operation. Such factors can be measured and utilized in selection, but in all work on safety of street railway operation it is apparent that selection techniques will reduce accidents only to a definite minimum, beyond which the study of accident susceptibility and, particularly, of emotional difficulties in individuals on the job becomes necessary if the optimum safety is reached.

EUROPEAN RAILWAYS

Attempts of certain Germans to apply Munsterberg's selection test as a safety technique led to the same conclusions and modifications that characterized the American experiment. A test developed by Boberg and Sachs (1920) illustrates this work. Their apparatus consists of a moving black belt representing the tracks, with holes in it representing pedestrians and vehicles. Each hole, when illuminated from below, requires attention by the operator; certain ones require response, others do not. The holes can be lighted unexpectedly at different distances from the operator, a selective response of controller, brake, or bell is called for, depending upon the distance and the nature of the signal given. Other distraction signals are also given. The test lasts about 20 minutes, which is long enough to introduce the element of fatigue. Results of this test in the elimination of accidents were, however, indecisive.

AMERICAN RAILROADS

"Safety First" seems to be the motto of many of the major railroads of the United States today, but originally scientific study of safety was slow to receive attention. The very high accident rate that characterized railroading in the years following the Civil War was due in large part to the character of the equipment used, particularly the brakes, draft gear, and signaling devices. As a consequence a great proportion of the effort to improve safety went into creating and installing new equipment. This step culminated in such installations as "dead-hand" automatic train-stopping equipment and experiments with radar. Only recently has there been a shift in effort to understand the human factors involved in safety. This situation has been hampered by rules, practices, and traditions which grew up at a time when the services of modern psychologists were not available. Hence psychological practices of the railroads are almost all a result of empirical experience, with little attention given to possible technical improvements developed in similar occupations.

The safety regulations of the Interstate Commerce Commission deal almost exclusively with materials and devices. For example, the Commission can and does regulate in detail the conditions under which signal apparatus can be operated, but it has no jurisdiction over the standards of selection, training, or the working conditions of signalmen who install, repair, and inspect signal apparatus. Rules relating to personnel, such as that limiting the working day to 16 hours, are rule-of-thumb propositions that were made when train speed and other significant factors were quite different from those now in operation. They have not been revised in the light of modern studies of fatigue or other relevant knowledge.

While the improvement in the operation of mechanical equipment reduced accidents remarkably, there were still unaccounted for those that were due to human failure. To eliminate these the Interstate Commerce Commission operates demonstration cars for the instruction of the men in efficient handling of equipment. The men are subjected to frequent examinations on the use of new devices. Surprise tests conducted while trains are in operation and surprise inspections in roundhouses and yards keep the men constantly alert to the dangers involved in rule violation. Whatever effects such scrutiny may have upon the personality of employees, the safety record of railroads in recent years is impressive testimony to the efficacy of the methods in use.

Safety campaigns similar to those in other industries, involving the incentive of competition, and the use of discussion groups to inform workmen of mechanical devices, have served further to reduce accidents on those roads with the best records. Individual attention by road foremen, trainmasters, and other specialists, has been given to accident-prone operators. The almost universal use of "brownies," or demerits, which are recorded against the personnel record of the individual has given alert management the opportunity to discover such individuals, but the punitive aspect of such a record probably intensifies rather than alleviates the emotional difficulties of the men. Workers commonly look upon their personal record as a sort of cumulative debit sheet on which, unless they are very careful, they will accumulate enough "brownies" to cause them to be discharged. On certain roads this situation is partly offset by awarding "safety buttons" or similar honors to men with a long accident-free record. There has been less use of the "brownies" system in the nonoperating fields with a proportionate decrease in the punitive attitude taken.

SELECTION

One of the consequences of studies in safety was the development of preselection tests. Selection for the purpose of promoting safety frequently resulted in improvement in other elements of efficient operation. Slocombe (1927) showed that efficiency in coasting, which reduced current consumption, and hence, costs, was closely correlated with an accident-free record. Preselection primarily to improve efficiency per se was, however, not characteristic of early efforts in America.

PRESELECTION IN EUROPE

The earliest applications of psychology to European railroads were intended to test all of the relevant factors possible. Lahy (1908) investigated energy consumption among transportation employees by means of the pneumograph, sphygmograph, plethysmograph, and electrocardiograph. In 1921 he established a battery of psychological tests in which he measured distributed attention, reaction time, motor suggestibility, emotionality, immediate memory, and perception of speed. Candidates for operating positions in Paris street cars were required to respond to street conditions as depicted on a motion picture screen by appropriate action in handling dummy controls of a street car. This record of reactions was compared with that of a successful operator in the same situation. A record of energy consumption was added to the scores on the test battery for prediction of successful operators. As a result of the use of this procedure the proportion of motormen dropped for incompetency was reduced from 20 to 3.4 per cent.

The Greater Berlin Tramway Company adopted apparatus to measure attention and response of operators which was quite similar to that used by Viteles. The selection of shopmen began in Dresden in 1921. By 1925 the whole operating staff of the tram company was being tested to indicate fears, control, vision, knowledge of regulation, and reaction in time of danger. Tramm (1924) stated that the reliability of these tests was from 85 to 90 per cent and Piorkowski (1922) reported that a reduction of 12 million marks was effected through their use in selection over a period of one year. Heidt (1925) modified these tests to include sensory capacity (e.g., sight and hearing range), attention, concentration, reaction time, power of decision, endurance, speed of work under pressure, agility, intelligence, judgment, memory, and conscientiousness. Panoramic tests, after Lahy, were used in the selection of railroad engineers. Puzzle boxes were introduced for testing work under difficult conditions, with interruptions and exhortations to speed up accompanying the test. Endurance was measured by the Dubois ergograph, blood pressure by the sphygmomanometer, intelligence by the Ebinghaus completion test, distributed attention by requiring concentration on a story being read while solving problems. Switchmen were tested on apparatus consisting of a number of moving objects, with intersecting paths which must be controlled to avoid collision. This test was said to reveal ability to judge speed and clearance. Seventy-five thousand men were tested in this manner, and Nestor (1934) reports that selection from these tests agreed with operating efficiency on the job in 90 per cent of the cases.

The testing program for the selection of railroad employees spread from France and Germany over most of Europe following World War I. Russia was using selection tests in transportation as well as in other branches of industry on a large scale by the later twenties. In 1938 the Transportation Section of the Leningrad Soviet employed 80 selection specialists including 15 physicians, 24 technical assistants, 14 psychotechnicians, and 27 other assistants. The results of the testing program, however, were largely nullified by the refusal of the authorities to make use of them. The refusal was based upon a social axiom of the Soviet philosophy which denied individual differences, and Viteles (1938) could find no evidence of the validity of the testing program.

AMERICAN RAILROADS

The empiricism that marked the safety program of American railroads likewise characterized the approach to personnel selection. With the exception of tests for sight and hearing the medical examination used in selection is not adapted particularly to railroading. Any standard color blindness test might be used. The Holgrem wools, which are commonly used for this purpose, require a finer discrimination than is needed, inasmuch as only three mixed colors, orange, purple, and green, are used in signaling, and these are standardized by suppliers of glass. None of the selective tests for reaction time, metabolism, or discrimination commonly found abroad are used by the major railroads. School records, character recommendations, and references by foremen or fellow-workers form the primary evidence in selection. Those who operate trains—firemen and brakemen—must make a number of student trips with other older and more experienced employees. During these trips those who fail to receive the recommendation of their superiors are dropped. Cottrell (1940) found that this recommendation is as likely to serve to maintain the mores of the working group as it is to measure prospective technical efficiency. The enormous turnover of railroad employees that resulted from the demands of the country during World War II challenged the railroads to improve their methods of selection, but the limited type of applicant available would undoubtedly have rendered the validity of a program of selection at least questionable. Whether for this or other reasons, we know of no large railroad system that altered its program of selection to include the tests found valuable in Europe and in other American industry.

TRAINING

Scientific psychology has seldom been applied to the training of the rank and file of railroad employees in the United States. After acceptance by enough of his fellows, the training of the student fireman or brakeman is almost entirely in the hands of the engineer or conductor with whom he works. With enough experience to pass the book of rules, he is qualified and presumably the equal of any other person with a similar classification. Seniority determines much more than efficiency the rating he will have insofar as it affects his pay check or the work he will do. Under the circumstances there is little that either the railroad or the individual can do to improve his operations. Shopmen are trained by a similar system of apprenticeship, though in this case the apprentice training may be fairly thorough and include the necessary technical knowledge and practice at the skills. A few railroads cooperate with a vocational program in the schools, but for the most part apprentice instructors teach the necessary technical knowledge, with the "school of hard knocks" constituting the primary instructor.

TRAINING OF SUPERVISORS

A notable exception to this method is found in the program of the Pennsylvania Railroad for testing and training the supervisory staff of three of its divisions (Hersey, 1941). In an investigation made in 1932 Hersey emphasized the emotional aspects of the railroad worker's life off the job, as well as on it, and their effects upon railroad operation. As a result he was able to make a thorough job analysis of the Pennsylvania supervisors. Supervisory work sometimes required a high proportion of concentrated reasoning; other work made use of meditative, relaxed reasoning; some of the work required rapid and accurate speed of perception; some of it required more time to discover and correct faults; while ability to understand and follow instructions was of extreme importance on other jobs.

Psychological tests were designed on the basis of this analysis with due weight being given to each of the factors. A rating scale with items that could be understood by railroad supervisors was evolved. A class was organized to teach the method of rating. The supervisors rated each other; they were also rated by their own superiors and by the immediate superiors of their superiors. This program was supplemented by a 45-minute interview between each supervisor and a psychologist. In these interviews guidance, leading to improved supervisory efficiency, was given.

This testing and rating program is used by the Pennsylvania Railroad in training and in rating supervisors for promotion. After five year's use it was reconsidered and the responsible officials considered that it had raised supervision to a high level of efficiency and the attitudes of workers toward supervisors was noticeably improved. The tests and ratings stimulate the supervisors to self-improvement and provide instruments of guidance to that end. The prediction of the combined battery of tests, ratings, and interviews was more than 96 per cent accurate.

MORALE

EMPLOYEE MORALE

Employee morale programs of the railroads are traditionally paternalistic. In consequence many of the employees have resented them and identified themselves much more fully with the union than with their employers. Some of the railroads have tried to develop unity among employees; notable among their attempts is the widely publicized "Old Timers Club" on the Union Pacific. This club has created a good deal of satisfaction among employees as well as provided good publicity for the company and its officers. During the twenties many of the companies introduced company-sponsored "independent" unions to replace the shopmen's unions, which were highly demoralized as a result of the strike in 1922. When the depression demonstrated the incapacity of these unions to resist the stringent cuts in employment which accompanied the depression, a great many of the men returned to the old-line unions, which have traditionally resisted all paternalistic devices intended to identify the worker with the company. In consequence, a good many efforts that seemed to be succeeding in the late twenties were dropped in the thirties. Union-Management cooperation which was most widely identified with the "B & O Plan" depended almost entirely upon suggestion and other empirical attempts to solve personnel problems, thus making little use of the body of techno-psychology developed in other industries.

PUBLIC MORALE

The railroads have made generous use of psychological advances in their institutional advertising. These campaigns have done a great deal to restore to the roads the public esteem in which they were once held, though much of the effect may have been offset by public experience with overcrowding, and what seemed to be widespread inefficiency and discourtesy during the war years. During the war great promises of a new and brighter day in the postwar world characterized railroad institutional advertising as it did that of many industries. Whether such romantic and sometimes fantastic promises will serve to attract or repel the public is not yet clear though some signs of a reaction against them have been visible. A kind of nostalgic romance on the part of the public toward the industry has been furthered by such devices as the displays which the industry presented at the World's Fair of 1940 in San Francisco and the Railroad Fair in Chicago in 1948. The Southern Pacific sponsored a series of children's clubs devoted to the perpetuation of the romance of the rails, and other roads took rail fans on good-will tours when they were given unusual insight into the operations of modern railroads.

MARKETING TRANSPORTATION

The railroads have fought a losing fight to maintain passenger traffic in competition with the automobile, the bus, and the airplane. Part of this loss has been due to the technological superiority of their competitors. Much has been due to the marketing practices of the railroads which have frequently violated the most elementary psychological principles. Railroads are slowly recovering from the public-be-damned policy of the days when they were a monopoly. A number of roads are training their personnel in salesmanship. Special courses in public speaking are given for traffic agents of the New York New Haven and Hartford Railroad. The Illinois Central has introduced general courses showing the relations of the railroads to other aspects of the society in which they operate. Until very recently a special director of courtesy and etiquette taught Union Pacific trainmen to meet the public graciously. While earlier the company assessed the results as favorable, abolition of the position makes them at least appear doubtful. The attacks on the industry by Mr. Young of the Chesapeake and Ohio Railroad shows that there is still much to be done before the full application of psychological principles in marketing railroad service will be possible.

The Union Pacific initiated marketing research in the industry through questionnaires circulated by its employees among the public to discover the causes of loss of traffic. On the basis of these complaints modification of policy was undertaken and the employees were encouraged to send in "Traffic Tips." This was part of the "Independent Union" program of the company and was rather generally abandoned when the older unions regained dominance.

The methods of recruitment, training, supervision, promotion, grievance procedure, and many other aspects of railroading are largely the result of the traditional social organization that grew up in the industry prior to the development of scientific psychology. The industry pays a penalty for its failure to adopt widely the tested results of this science. The costs will probably have to be much more apparent to both management and labor in the industry, through competition with those both within it and without, before psychotechnology will receive the attention which its record in other American industries indicate that it is entitled to receive.

READING REFERENCES

Cottrell, W. F., *The Railroader*. Palo Alto, Cal.: Stanford University Press, 1940.

Hersey, R. B., "Experimental Application of Psychological Tests and Merit Ratings to Supervisory Selection, Advancement and Training on the Pennsylvania Railroad." *Studies in Personnel Policy*, National Industrial Conference Board, 1941, No. 32, 61–69.

Nestor, I. M., "Vocational Tests on the European Railways." *Human Factor*, 1933, **7**, 11–23; 51–58.

Viteles, M. S., *Industrial Psychology*. New York: W. W. Norton & Company, Inc., 1932, Chap. 14, "Tests in the Transportation Industry"; Chap. 18, "Accidents in the Transportation Industry."

———, "Industrial Psychology in Russia." *Occupational Psychology*, 1938, **2**, 85–103.

52. AVIATION

JOHN C. FLANAGAN, PH.D.

Professor of Psychology, University of Pittsburgh; President, American Institute for Research. Formerly: Director, Aviation Psychology Program and Chief, Psychological Branch, Office of the Air Surgeon, Hq. AAF; Associate Director, Cooperative Test Service, American Council on Education; Lecturer, Tests and Measurements, Teachers College, Columbia University.

INTRODUCTION

Psychologists began to study problems of aviation intensively during World War I. Almost none of the group participating in this research, however, continued their activity in this field after the war was over. During the period from 1920 to 1930 practically the only work done in this field was at the School of Aviation Medicine of the Army Air Forces at San Antonio, Texas. Here Dr. L. J. O'Rourke served as a temporary consultant and developed a complex coordination apparatus which after further modification was experimentally validated and became one of the mainstays of the Aircrew Classification Test Battery in the AAF in 1942. Dr. Ross A. McFarland was one of the first to become interested in aviation psychology prior to World War II. His early publications related mainly to the effects of high altitude.

The revival of interest in aviation psychology in this country was somewhat slower than in other countries and was led not by the military organizations but by the Civil Aeronautics Administration, Department of Commerce. Dr. Dean R. Brimhall, Director of Research for that organization, obtained funds in 1939 for psychological research in the selection and training of civilian aircraft pilots. These funds were administered by the National Research Council through a committee established for this purpose called the Committee on Selection and Training of Aircraft Pilots. In 1941 and 1942 the Army and Navy established projects in aviation psychology.

This discussion is primarily based on the recent findings of these research programs.

SELECTION OF AIRPLANE PILOTS

MILITARY PILOTS

The war provided an opportunity to try out a wide variety of tests on an unusually large sample of candidates for pilot training. These findings are reported fairly completely in available publications. A few generalizations may be made. A major result was the substantial validity found for several apparatus tests. These included such functions as complex coordination (using stick and rudder), two hand coordination, discrimination reaction time, and rudder control. Tests of this type found not to be of value included steadiness, finger dexterity, dotting, tapping, and simple reaction time.

Various conventional printed tests dealing with abstract symbols which had been used successfully in educational institutions were given a thorough tryout. Tests of vocabulary, arithmetic computations, and memory were found to have little predictive value. Tests of arithmetic reasoning and reading comprehension had a small amount of predictive value.

Among various types of printed tests in which the task was one of visualizing or identifying objects from symbolic or pictorial representations a number of predictive tests were found. These included tests of visualizing movements, identifying photographs, reading instruments, and comprehending the operation of mechanical devices.

Although none of the various stress tests used was shown to be effective, two types of tests relating to the applicant's personal background and experiences were found to be of value. One of these asked for biographical data which was evaluated in terms of an empirical scoring key derived from the follow-up, on several thousand student pilots. The second test, instead of asking which activities had been engaged in at some previous period of time, raised specific questions of information which were constructed so that only a participant would be likely to answer them correctly.

Combinations of scores from these various types of tests gave predictions of great practical value. In the Army Air Forces, for example, the validity of the combined aptitude score (pilot stanine) was found to be 0.66 in a group of more than a thousand men who were accepted for training regardless of their scores on the tests. The multiple correlation in this sample was 0.69. The correlations are biserial correlations between the scores and the criterion of graduation or elimination from pilot training schools. Figure 1 shows the experience in the AAF regarding the relation between pilot stanines and success in primary training schools. During the last year of World War II only men with the highest three stanines (7, 8, and 9) were accepted for training for positions as aircrew officers.

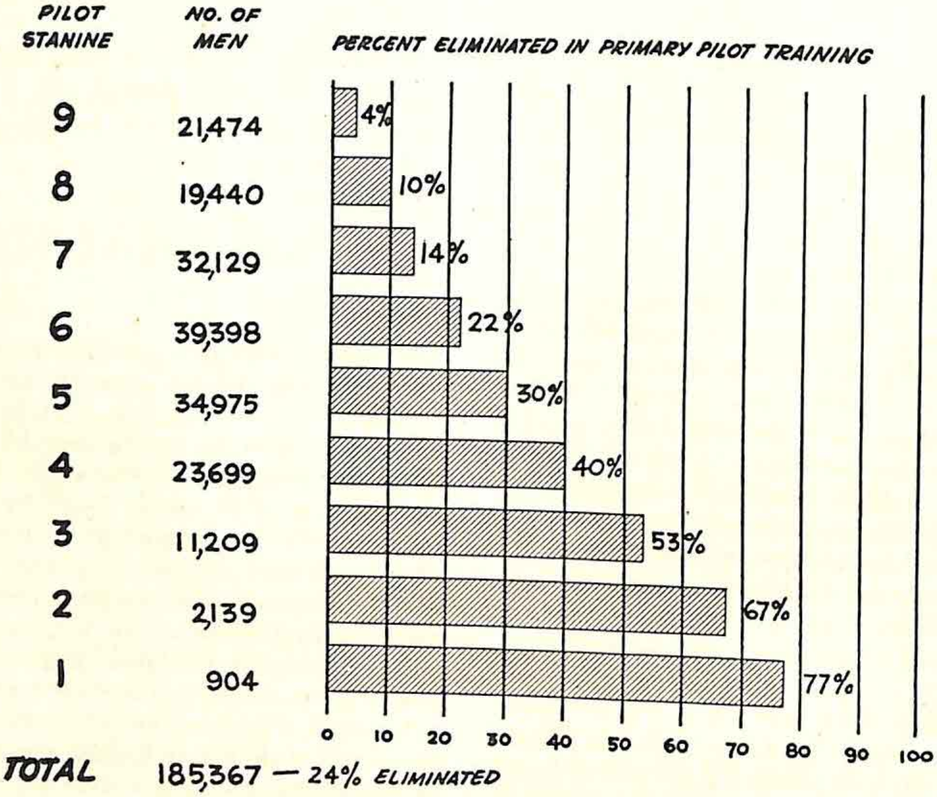


Figure 1 (VII-52): *Army Air Forces Aptitude Test*

The bars indicate the proportions eliminated at each pilot stanine. Elimination was for flying deficiency, fear, and own request. Flying experience credit is included in the stanine score. The data are from classes 43–F through 45–H. Men with low stanine scores were disqualified for training in the later classes; most of the men with low stanines included in the chart entered primary schools early in 1943.

In the summer of 1947 the Aviation Psychology Branch, Bureau of Medicine and Surgery, US Navy, with the cooperation of the Aviation Psychological Program of the Army Air Forces, initiated an intensive study of personnel in preflight pilot training at the Naval Base at Pensacola to determine the relative validity of a number of new tests, many of which were developed but not adequately evaluated during the war period.

The emphasis in selecting pilots during the war was on obtaining men who were, or could learn to be, skillful flyers. Since the war emphasis has shifted to give more weight to the factors necessary for superior commanders.

PRIVATE PILOTS

Light planes can now be flown by almost anyone who can drive an automobile. However, younger individuals learn much more rapidly and easily than older persons. Recent studies sponsored by the Civil Aeronautics Administration, and carried out under the supervision of the National Research Council at the Ohio State University, indicate that even a large proportion of individuals having very defective vision in both eyes or having only monocular vision are able to complete their training and pass the flying examination for the Private Pilot's License.

In connection with the Civilian Pilot Training Program of the CAA in 1942–43, the Committee on Selection and Training of Aircraft Pilots of the National Research Council examined 67,067 candidates for pilot training. The three tests used for selection included an Inventory of Personal Data, a Test of Mental Alertness, and a Test of Mechanical Comprehension, which were administered in several hundred centers. About 30 per cent of the applicants were rejected on the basis of the test scores. Biserial and product moment correlation coefficients obtained between test scores and the criteria of pass-fail and flight grades for a sample of about four thousand individuals passing the tests tended to be low. The highest validities were obtained for the Test of Mechanical Comprehension and a test of Aviation Information which was not used as a basis for selection.

AIR TRANSPORT PILOTS

The air transport pilot has only recently been the subject of extensive investigation. In the summer of 1946 a study was made by the American Institute for Research of four hundred first officers for one of the major airlines. This study demonstrated a significant degree of relationship between adaptations of several tests used in the Army Air Forces Aircrew Classification Test Battery and the average of a number of ratings of pilot fitness given these men in the period of between 6 and 12 months during which they had observed their work. The tests showing the greatest relationship with this criterion involved speed of reading instrument dials, knowledge of aviation, discrimination reaction time, and visualization of maneuvers. The American Institute for Research and the Committee on Selection and Training of Aircraft Pilots of the National Research Council are conducting further studies of the requirements of airline transport pilots at the request of the Civil Aeronautics Administration.

OTHER AIRCREW

The long-range heavy bomber type of operations during the war made it necessary to add a number of additional crew members. The duties of certain of these, particularly the navigator, radar observer, and flight engineer appear likely to become a definite part of long-range air transport operations in civil as well as military flying.

All of these positions were the subject of intense study during the war. However, the latter two positions did not develop until late in the war. Therefore only in the case of the navigator was there an opportunity for extensive follow-up of selection test results. It was found that to an even greater extent than the pilot the navigator needed the abilities to read instruments accurately, to use maps and photographs, to understand verbal materials, and to solve arithmetic reasoning problems. Also, whereas tests of vocabulary and arithmetic computation had been found to have no predictive value for pilots, they had substantial validity for selecting navigators. The predictive value of the navigator stanine was found to be between 0.60 and 0.70 in terms of the bi-serial correlation coefficient for this score with graduation-elimination from navigation training schools. Since the reliability coefficient for the training school criterion of graduation-elimination appeared to be between 0.50 and 0.70, the maximum obtainable coefficient for large samples with these criteria would be between 0.71 and 0.84.

In this instance, as with the pilot training schools, careful study indicated that success or failure in the training schools was based on the systematic judgments of a number of instructors and examiners with respect to abilities which were important for success in operational work of this type. Therefore, although the criterion used in most validity studies was not an ultimate one, its relevance for selecting appropriate personnel appeared to be satisfactory.

In general summary it may be said that extensive research has been carried on with respect to the selection of various types of pilots and other aircrew members, especially in military operations. Particularly effective in selection for success in training have been the pilots' and navigators' stanines. Although this work provides an excellent basis for present practice, the rapid changes in aviation equipment and procedures can be expected to result in major changes in requirements and a continuing program of research is essential to maintain satisfactory selection procedures.

TRAINING AND EVALUATION OF PROFICIENCY

CONTENT AND TECHNIQUES OF INSTRUCTION

The recent contributions to problems of the training and evaluation of the proficiency of aircrew personnel are primarily in the area of techniques of study rather than finished products. The importance of the content of instruction has been emphasized by findings during the war. Too often hastily established schools were later found to be teaching irrelevant and sometimes even detrimental skills and information.

A basic method of study which has been found very effective is the determination of the *critical requirements* for success in a particular activity. A *critical requirement* is defined as one which has been found to have made the difference between success and failure in a substantial number of instances.

Studies of combat operations and interviews with combat personnel were carried out to determine necessary or desirable changes in the content and methods of flight training. Similar studies are now being made of the *critical requirements* of airline transport flying.

An important research technique which led to notable improvements in flight training was the development under the auspices of the Committee on Selection and Training of Aircraft Pilots of procedures for recording flight instructors' remarks. These recordings led to the development of (1) Patter for Elementary Flight Maneuvers and (2) Fundamentals of Basic Flight Maneuvers. *Patter* provides a model presentation to be given by the instructor in the air of each of the elementary flight maneuvers. This material was based on careful analysis of instructional practice and with due regard for the psychological principles of learning. *Fundamentals* outlined the basic facts which should be learned before going into the air for instruction on a specific maneuver.

This committee also developed a 30-hour course on the theory and technique of flight instruction. In 1947 they prepared a booklet for publication by the CAA including special teaching techniques which individual instructors had found effective. These instructional "tricks" were reviewed by a number of instructors before publication.

INSTRUCTORS

As indicated in the discussion above, more than the ability to fly is essential to being a good instructor. Almost invariably, however, flying ability has been the major if not the sole measure used in selecting instructors. It was found very difficult to evaluate instructors. Observation of actual instruction or practice teaching, and systematic rating and recording of specific aspects of such instruction were found useful. Another technique which was utilized extensively was obtaining ratings of instructor behavior from students.

Although conclusive results are not available, results from a number of studies suggest that in addition to flying ability the instructor should have (1) insight into flying problems and the ability to express these ideas clearly, (2) an understanding of the individual and of learning principles and techniques, and (3) suitable temperament and attitudes.

CHECK RIDES AND STANDARD FLIGHTS

Methods of evaluating the proficiency of pilots have not yet reached a satisfactory level of reliability and validity. Much progress has been made in recent years, however. The Committee on Selection and Training of Aircraft Pilots took this as one of its major problems and standard flights, rating scales, standardized inventories for recording observations of pilot performance, graphic flight recording devices, and techniques for the photographic recording of flight and control movements were developed. All of these procedures had specific advantages. The graphic and photographic recording techniques produce objective records descriptive of actual flight behavior, but they are very expensive to analyze. The standardized inventory has proved useful but although the recording of observed behavior is quite objective, the precise nature of the task and the relative importance of the various components are usually not clearly defined to the pilot being tested. For this reason men with the same ability may make very different scores by giving a greater proportion of their attention to doing good work on one portion of the task at the expense of other aspects. For example, in making a turn the pilot can either emphasize the avoidance of slips and skids and tolerate more variation in altitude or he can hold the altitude as constant as possible and give less attention to the coordination of bank and heading.

In addition to the problems of objective recording of results and the objective interpretation of the nature of the task it has been found that it is difficult to control the nature and difficulty of the job itself. Planes vary widely in their flying characteristics and the calibration of their instruments. Wind, turbulence, density, and visibility also affect the task.

Findings from various studies suggest that the agreement between the ratings given on two independent check flights is typically represented by a coefficient of correlation of about 0.20. In a typical study in the Army Air Forces the correlation coefficient for grades in primary training schools with grades in basic training schools was 0.27 for a sample of about three thousand pilots. A comparison of grades on the CAA certification examination for about a thousand pilots on the Private Pilot, Commercial Pilot, and Airline Transport Pilot Examinations indicated that the agreement between them was very close to zero when expressed as a correlation coefficient.

OBJECTIVE FLIGHT CHECKS

During the last year of the war the aviation psychologists working on pilot training projects in the AAF made the development of an objective scale of flying skill their main task. In studying the type of observations which could be made objectively it was found that good agreement could be obtained as to whether the air speed increased or decreased by more than a fixed amount over a specified period of time. Similarly, agreement could be obtained on most other specified instrument readings. The extent of bounce on a landing, the location of landing, and the degree of bank in a turn were also items which could be reported objectively. On the other hand, items, such as "looked around 50 per cent of the time," "satisfactory ground pattern," and "excessive use of the throttle," were found to be difficult to obtain agreement on.

To insure that all pilots set themselves the same task and were fully aware of the effect of their behavior on their scores, the scoring procedure on each item was explained to them in advance. For example, on a steep turn the pilot was told that he would lose one point for each 20 feet gained or lost from the starting altitude during the turn. He was also informed that his score would be zero on the maneuver if he failed either to reach a 60-degree bank in the turn as indicated by the cabane strut being parallel with the horizon or if his final roll-out was not within 45 degrees of the initial heading.

Studies of standardized maneuvers of this type emphasized the extent of variations in the condition of the plane and other equipment and in weather conditions. It was found that the scores made by a particular pilot on certain types of objective items repeated on the same day in the same plane agreed fairly well but that if the item was repeated on the following day in a different plane the extent of agreement became unsatisfactorily low.

Fortunately this was not true of all types of measures. An objective check-ride of single-engine instrument flying skill developed in the summer of 1945 showed a substantial amount of agreement between scores for the same standard check ride given on successive days by different check pilots. The correlation between these scores for a small sample of fifty-five student pilots was 0.46. The correlation between the sum of the total scores for the two check rides and an over-all rating of instrument flying proficiency made immediately before these rides by the student's instructor was 0.51 for the same sample. These figures suggest that it is feasible to develop a reliable and valid objective scale of instrument flying skill.

DESIGN OF EQUIPMENT

PERCEPTION OF NECESSARY ACTION

In selecting, training, and evaluating personnel who are operating complex equipment, the psychologist is frequently impressed with the idea that the task of the individual has been unnecessarily complicated by arbitrary factors in the design. This observation during World War II led to the assignment of several groups of psychologists in the various services and in civilian research organizations to make studies of factors affecting human efficiency so that these could be included along with engineering factors in planning and designing new equipment.

An example of the type of studies carried out in this field is the evaluation of aircraft attitude indicators on the basis of Link instrument ground trainer performance. This study was made in the Department of Psychology in the AAF School of Aviation Medicine. The findings indicated that the most suitable type of flight attitude instrument presented the pilot a miniature airplane on which he could observe directly the effects of his control movements and the resultant attitude of the plane. With this instrument the pilot had only to operate his controls in such a way that the miniature plane was directly perceived to be performing the required maneuver. The types of mental translations and allowances required by the attitude indicators commonly used during the war were entirely eliminated.

Another example of this type of study is provided by recent research in the Psychology Branch, AAF Aero-Medical Laboratory. Interviews with pilots revealed that a large number of them had been placed in dangerous situations by misreading the altimeter. Usually the error was in terms of one or more thousand feet. Examination of the dial led to the design of a laboratory experiment in which a number of new types of instruments were compared with the form in current use. A new indicator resulted, which practically eliminated errors of the type that had been the source of serious danger. This instrument was read more rapidly and accurately than the former altimeter even by pilots with years of experience using the old type.

Another study carried out by the Psychological Branch at Wright Field revealed that a fairly significant period of time was required to perceive the attitude of a plane and initiate corrective control action. Twenty AAF pilots were given a 3-minute disorientation procedure wearing goggles that transmitted light but permitted no pattern vision. The goggles were then opened and the times to initiate corrective control action and to complete recovery to straight and level flight were recorded photographically. For contact flight 1.35 seconds were required to perceive the attitude of the airplane in relation to the horizon and initiate corrective action. For instrument flight 1.55 seconds were required to perceive the attitude of the airplane from instrument readings and initiate corrective action. Recovery to straight and level flight was more rapid under contact than under instrument conditions.

EFFECTIVENESS OF CARRYING OUT REQUIRED CONTROL MOVEMENTS

In addition to the pilot's perceiving the necessary action it is essential that his movement of the controls accomplish the intended action. A series of studies was made in 1946 by the Psychology Branch, Aero Medical Laboratory regarding the relation of various design factors to efficiency in the operation of the pilot's controls. The findings from these experiments using a simplified cockpit mock-up are as follows: (*a*) control with the hands and arms was considerably better than with the feet and legs; (*b*) fore and aft (elevator) movements were slightly more efficient than lateral and rotary (aileron) movements; (*c*) comparable movements were made with approximately equal efficiency with either the wheel or stick; (*d*) tests under three conditions of average leg or arm angle (105°, 120°, and 135° at knee and elbow) showed virtually no differences in efficiency of control with rudder pedal or wheel (with fore and aft movement).

Another experiment was carried out to determine the direction of control knob movement which seemed more natural to pilots. It was found that when the indicator and control were in the same plane and that plane faced the operator, pilots consistently responded by rotating the control knob so that the part adjacent to the indicator was moved in the same direction as it was desired the indicator move. When the indicator and control were not in the same plane the pilot's responses were not consistent.

Work on these types of problems is only beginning. Results to date suggest that the psychologist can be expected to play a prominent role in equipment development in the future.

OPERATIONS

AIR SICKNESS

One of the general problems of aviation operations is air sickness. Recent studies in this field have been concerned with susceptibility, the analysis of precipitating causes, and the evaluation of proposed preventative measures. Although every individual is susceptible to air sickness there are large individual variations in the amount of motion required to produce symptoms. Statistics tabulated by the AAF School of Aviation Medicine indicate that between 10 and 15 per cent of all flying trainees become airsick at least once during the training period. However, only 5 per cent become airsick more than once. Approximately one per cent of the aviation cadets were eliminated from flying training because of chronic air sickness.

Various studies have shown that children are more susceptible to air sickness than adults. Apprehension, suggestion, fatigue, poor physical condition, and reading have been most frequently cited as conditions increasing susceptibility. The precipitating cause appears to be primarily the up-and-down motion of the plane rather than other types of motions.

It has been found that daily exposures to up-and-down motion as, for example, in a swing, decrease susceptibility to motion sickness. This confirms the findings with regard to aviation cadets indicating that the individual becomes adapted to the motion of a plane. Extensive investigation has been made of the effect of various drugs on the incidence of air sickness. Few drugs have been found which have any influence at all on susceptibility. The most effective of the drugs tried is hyoscine hydrobromide. A single dose of 0.75 milligrams is ordinarily administered one hour before flight. Caution should be observed in its use since larger doses will affect vision and inhibit sweating. This drug is effective only with a minority of persons. Preliminary experiments with a new drug, dramamine, indicate striking results both for prevention of air sickness and relief of its symptoms.

FATIGUE

As in other fields it has been difficult to obtain definite results regarding fatigue in aviation. Evidence from studies of fighter pilots indicates that these men can sit in one position in a small cockpit with no opportunity to make any large shift of position and fly their plane for 10 hours without any observable impairment in their skill at the end of this period. On landing they report they feel fine though perhaps a little stiff and sore, but the next morning they usually report that they feel tired.

Stress due to equipment failure, bad weather, or enemy action is reported by them to increase their feeling of fatigue. Individuals who have flown for long periods on several successive days report that it has taken them several weeks or even months to feel fully recovered.

Fatigue is a very complex problem and many factors other than duration of a specific flight make important contributions to the alertness and efficiency of the pilot. Much additional research is needed on this topic.

ACCIDENTS

Aviation accidents are very dramatic and are extensively reported in the popular press. The safety record of the commercial airlines has been good but the public is somewhat misled by the attention given to the occasional crash of a large passenger plane with the loss of a number of lives and does not get a true perspective on the relative safety of the airways as compared with other modes of transportation. The number of fatalities per hundred million passenger miles flown was 1.63 for the airlines as compared with 0.25 for the railroads for the period 1943–46.

Most of the accidents which do occur are connected with take-off or landing. In the accidents investigated in 1944 nearly half occurred on landing and almost another quarter on take-off. An analysis of 257 fatal accidents revealed that 65 per cent involved stalls. Recent research indicates that even experienced instructors do not recognize stalls accurately and a number of stall warning devices have been developed which sound alarms and flash red lights when the aircraft is approaching a stall.

In a recent study carried out for the CAA by the American Institute for Research under the auspices of the National Research Council, airline pilots reported that two factors comprise a large proportion of the dangerous situations encountered, namely ice forming on the wings and cross wind on landing.

Research in the Army Air Forces demonstrated that both age and aptitude, as shown by psychological tests, are related to accidents. On the basis of a study of 4,638 accidents occurring in the AAF in 1944 and 1945 it was found that the highest accident rates were for pilots under 22 years of age. This was still true when the amount of experience in terms of hours of flying was controlled. The lowest accident rate were among the age group 26-34. The increase in accident rate beyond the age of thirty-five was more marked for fighter planes than for heavy bombers. The types of errors made more frequently by the pilots with long experience were "failure to observe" and "misuse of controls."

Various studies of aptitude in terms of pilot stanines showed that the accident rate was two or three times as great for the low-stanine as for the high-stanine group.

ALL-WEATHER FLIGHT

The latest development is in new procedures to make possible all-weather flight. The present radio range and instrument flying procedures make it possible to fly in overcast weather. The problem of landing when the airfield is completely covered by low clouds or fog has not, however, been satisfactorily solved.

Psychological research is in progress with respect to several of the problems related to blind landings, such as the most easily learned signals. The relative speeds of perception and response for various types of signals and also the relative frequency of occurrence of errors are being studied.

The aircraft engineers foresee the day of fully automatic control from a central point with the plane flying on an unseen but definitely chartered track with the pilot merely riding along to see that the equipment is functioning properly. In the meantime the utilization of psychological principles and techniques will be of great assistance to both military and civilian aviation.

READING REFERENCES

Army Air Forces, Aviation Psychology Program, *Research Reports,* Nos. 1–19. Washington, D.C.: Superintendent of Documents, 1947.

Civil Aeronautics Administration, Division of Research, *Report Series,* Nos. 1–70. Washington, D.C.: The Administration, 1942–1947.

Flanagan, John C., "The Experimental Evaluation of a Selection Procedure." *Educational and Psychological Measurement,* 1946, **6**, 445–466.

McFarland, Ross A., *Human Factors in Air Transport Design.* New York: McGraw-Hill Book Company, Inc., 1946.

CHAPTER VIII

BUSINESS PSYCHOLOGY

53. Product Testing	Philip G. Corby
54. Copy Testing	Albert D. Freiberg
55. Advertising Media Research	Darrell Blaine Lucas
56. Customer Research	John G. Jenkins
57. Radio and Television Audience Research	Ernest Dichter

53. PRODUCT TESTING

PHILIP G. CORBY, M.A.

Associate Director of Market Research, Psychological Corporation.

Product testing through determining consumer reactions to new or improved products is as old as invention and much older than psychology as a science. The inventor of a new product, or of an improvement upon an old one, has usually made use of his family and his friends to try out or test his product. In many laboratories new products are first tried out on the laboratory technicians themselves, their friends, their families, and workers in various parts of the business. This first test passed, the public in one way or another becomes the subjects for further testing, usually through sales in small market areas.

Marketing has developed with civilization until under present competitive conditions the cost of marketing a new or improved product is very high indeed, even if done on a limited scale. It is, therefore, important that new products be pretested under carefully controlled and unbiased conditions such as can be achieved through the application of experimental techniques which several decades of work in psychological laboratories have developed. Scientific product testing is the practical application of the experimental methods and techniques of the psychologist to the measurement of consumers' discriminations and preferences. This is accomplished through laboratory and field studies.

The manufacturer who has planned an improvement in his product is vitally concerned with the extent to which his customers accept the improvement and even more so with the extent of his competitors' customers' acceptance of the new product. Is the acceptance sufficient to justify any change in the cost to the consumer? Will the change add new uses for the product? Will consumers recognize the new uses readily or will extensive advertising be required to inform consumers of the value of his changes? When the taste or flavor of a product is affected, the problem is complicated because while the immediate reaction may be favorable, continued use may be less appealing. Safety razors introduce the problem of approval of the shaving characteristics over an appreciable period of time. A change in a razor may initially appear very desirable but be appreciably less so after the growth of the beard has adapted to the cutting edge of the new razor. The razor may introduce changes in the way it is used by its weight distribution, size of handle and other characteristics. These are some of the types of problems which should be answered before placing a new product on sale in even a limited or test market.

This variety of problem requires a very carefully controlled testing procedure to determine consumer preferences. In some cases the controls required indicate the need of laboratory study. In other situations the laboratory imposes conditions so artificial that the study of consumers' reactions must be in the normal place of consumption for the trial of the new product. For the manufacturer, the first consideration is the elimination of bias in a consumer test of his product. Preliminary work with employees or interested groups may be biased. In order to obtain an unbiased result it is necessary to apply the controls of psychophysical measurement familiar to psychologists. These principles must also be applied to field studies where the test must be brought to the consumer's home or other normal place of consumption.

The laboratory frequently permits more rigid control and fewer subjects may be used. The field test situation permits less control and a larger number of subjects is usually necessary in order that variables not under control may be analyzed statistically. The normal variations in a consumer's use of a product may frequently be critical when a product is being changed. For example, a change in a cereal product could be preferable for some consumers whose cooking habits are of one sort and be unsatisfactory to others who use different cooking methods. These problems will be illustrated by descriptions of five product tests, two conducted in the laboratory and three conducted in the field.

LABORATORY STUDIES

CRISPNESS OF CEREALS

The most desirable package for a flake cereal, for example, corn flakes, is one which is relatively inexpensive and brings the product to the consumer in a dry, crisp condition ready to eat without heating. Crispness produces tactile sensations which are a part of the gustatory experience considered pleasing by the consumer. A manufacturer had designed a new package which his chemist had shown to be highly efficient in preventing moisture from robbing the cereal of its crispness. The problem of the manufacturer at this point was to determine: (1) whether or not the difference in moisture content of the cereal was sufficiently great for consumers to discriminate when the cereal was taken from the new and from the old package and (2) whether or not this new package was superior to competitors' packages.

Cereals were stored in a humidifier in the laboratory both in the new and old packages. Cereals of other manufacturers were similarly stored in their original package and after being repacked in the new package. The effect of the new package could then be tested for several cereals to be sure that the discrimination of crispness could be attributed to the package. They were kept several days before use in the test under extreme humidity unfavorable to the maintenance of dryness and crispness.

Thirty-six trained subjects were used, each participating in an extensive series of trials. They were presented in each trial with a pair of samples of the same cereal in small paper cups. Each pair contained the same cereal from the new package and from the original package for that cereal. The subjects picked up with their fingers in the order directed a small sample of each cereal to taste and report which of the samples was the crisper. Whatever visual cues or tactile cues existed in the picking up of a sample would have been a part of the normal experience of the consumer in eating the cereal and were allowed to operate without control.

The results showed that the subjects readily discriminated between the crispness of cereal from the new package and either the same cereal from the old package or a competitor's package. Preference reports also obtained from the subjects were clearly in favor of the new package.

TACTILE PERCEPTION OF FABRICS

The introduction of rayon yarns created many manufacturing problems in materials for women's clothing. The earliest rayon materials were made up in a manner similar to the weaving and finishing of older silk materials. Manufacturing developments produced a wider variety of materials in which rayon could be used but the nature of rayon fiber produced in certain weaves a different reaction when the material was handled. A manufacturer of rayon yarn had been experimenting with a variety of materials in which his yarn could be used in competition with other materials. A certain amount of consumer resistance had already been met because the "feel" of a rayon material was not familiar to many women. What would be the consumer's reaction to materials that had been made up experimentally but not yet marketed?

Most important for experimentation were considered the tactile sensations in handling rayon as compared to similar weave materials produced from silk, cotton, or wool. Conditions were set up under which the subject's report of the tactile sensations and their relative pleasantness or unpleasantness could be made both with and without visual clues. The subject was asked to name the type of fiber from which the sample material she handled was made, whether it was rayon, silk, cotton, or wool, because at the time of this study rayon materials were considered by many women to be inferior or undesirable. The results of this study showed which rayon material produced unpleasant tactile sensation and which did not. It also showed the extent to which women could discriminate between material made from rayon yarn or from other yarns. From the results of this study the manufacturer could determine which types of materials from rayon yarns could be promoted with weavers, finishers, and designers with the least consumer resistance because they were made of rayon.

This study was conducted for the manufacturer of rayon yarn who wanted to convince the weavers, finishers, and dress manufacturers of the desirability of wider use of fabrics made from rayon yarns and it has been given a detailed published report because its purpose was to promote an industry (Lazarsfeld and Ripin, 1937).

A considerable number of studies have been made by psychologists similar to the two just described. Many remain unreported in the psychological literature because a manufacturer in a competitive economy does not believe that his competitor should necessarily profit by what he discovers about his product and his competitors' products. Similar studies have appeared in the psychological literature but have usually been written and titled from the point of view of psychological implications and not of practical application.

FIELD STUDIES

It is important in merchandising many products to have: (1) attractive colors for display purposes, which will permit the product to be readily observable against the background of competing or adjacent products, and (2) colors which harmonize with the product. The following field study illustrates a color preference problem. The natural conditions at the time of purchase were used instead of a laboratory approach.

COLOR COMBINATIONS ON CAKE BOXES

A baker was dissatisfied with his package for a sponge layer cake sold to housewives to frost or use in a shortcake with berries or other fruit. Several articles usually found on the same shelves had been given new package designs and colors, leaving the sponge cake at a disadvantage in display. Package designers prepared a basic design with an open center through which the cake was visible and with several alternative color combinations all having relatively good display advantages. Some colors adjacent to the cake would be less attractive to the prospective purchaser; particularly the colors that might not suggest its use. The combination that would be most pleasing to the housewives was the question to be settled by tests.

To bring housewives who buy sponge layer cakes to a laboratory would have been a costly procedure because of the difficulty in locating them and giving them a mental set or attitude favorable to testing cake packages. To make the grocery or delicatessen the laboratory and utilize as subjects only those housewives who came into the store to buy such a cake, sacrificed a few controls available in the laboratory but guaranteed a highly favorable mindset among the subjects. While this procedure was expensive it was appreciably less costly than getting proper subjects into a laboratory. The use of unselected adults as subjects would have required a comparison with subjects selected according to habits of purchasing this kind of cake.

Fresh cakes in five new boxes were utilized to obtain color preference reactions, using both the familiar paired comparison method and rank order procedure. The experimenter operated in six widely scattered grocery stores and questioned a total of one hundred and eighty-two housewives of varying ages. The results were compared for reliability by the two methods. Two boxes with color combinations having one color in common were preferred about equally, and the baker and his advertising counsel had little difficulty in determining which of the two had the superior display advantage.

PREFERENCE FOR RECTANGULAR AND OVAL SPONGES

The manufacturer of a rectangular shaped cellulose sponge wanted to measure the preference for a new oval shape as compared with his present rectangular shape. His problem was: Would consumer preference be sufficient to justify marketing nationally one shape in preference to the other, or should he market both? Sponges vary in size and usually are used in the home for different purposes depending upon the size. The effect of shape upon the major use to which various sizes were put introduced certain complications in the field study that was planned.

EXPERIMENTAL PLAN

In a typical city one hundred housewives already adapted to the rectangular sponge and one hundred housewives unfamiliar with cellulose sponges were provided with sponges and instructions as follows:

	100 Housewives Unfamiliar with Cellulose Sponges	*100 Housewives Familiar with Cellulose Sponges*
Sample Sponges provided	Oval in two sizes to match present rectangular sponge sizes	Oval and rectangular shaped sponges in two sizes: medium and large
Instructions for using	Same as present sponge use	Instructed in most frequent uses for each size

The experienced housewives were located in upper and lower middle-class neighborhoods from a previous study in which customers accustomed to using sponges had been interviewed about their experiences with the rectangular sponge. The inexperienced housewives were similarly located by interviewing housewives until one hundred had been found who used natural sponges. A few days after the sponges had been left for trial, telephone calls were made to inquire about the extent of use and to suggest the use of sizes or shapes not yet tried.

Four weeks after leaving the sponges, field workers called on the housewives to interview them and determine: (1) if the sponges had been used, (2) the uses to which each size and shape had been put, (3) the preferences for the size and use of each shape, and (4) the housewives' reasons or basis for their preferences. Approximately 85 per cent of the housewives had cooperated by trying the sponges under normal conditions of use.

THE RESULTS

A slight majority of the experienced housewives showed a preference for the oval shape. This held true for most of the uses to which the sponges of different sizes had been put. However, a slight majority of the housewives who had not previously used a cellulose sponge preferred the rectangular shape for most uses. But both shapes satisfied the housewives and little would be gained by having both available. The decision to market one or the other, or both, was entirely a matter of manufacturing and distribution costs, and the merchandising and advertising involved.

TEST PREFERENCES FOR TWO CEREALS

A manufacturer of breakfast cereal of a type cooked to be eaten hot, usually with milk or cream and sugar, was dissatisfied with his sales. His analysis of his problem left the following possible reasons as paramount: (1) size of package and price unfavorable compared with products of certain competitors, (2) inadequate, or possibly ineffective, advertising, and (3) changing consumer preference, leaving his product with a dwindling market. If consumer preference for his product is declining, changes in pricing and advertising can have only limited effectiveness in modifying the downward trend in sales.

To test consumer preference, packages of this manufacturer's cereal and of a principal competitor's were prepared under an unknown name and marked for use in serial order. One hundred and twenty housewives who used cooked cereals and were willing to cooperate in the test were selected as the subjects. The order of use was clearly arranged so that Package A was used first by one housewife and Package B by the next. Each housewife was instructed to cook the cereal in the same way that she cooked her regular brand and serve it to her family at the time that she would ordinarily serve that cereal. A preliminary interview with the consumers cooperating in the study had brought out the method of cooking in terms of amount of water or milk and cereal, type of pan used, and length of cooking. Each one of these variables affects both flavor and consistency in the cooked cereal and the sample was treated statistically for the influence of each of them.

The housewives were interviewed after their families had been served twice with each cereal and it was found that preference for the manufacturer's product exceeded that of his competitor's. Since his sales were considerably lower than the competitor's he set out to revise his package size and price to compete on a more favorable basis. Also, there was a correlation between preference and cooking methods which lead to changing the cooking instructions printed on the package.

PSYCHOLOGICAL SERVICES

The first three studies above require the services of a psychologist to be certain of sufficient training in experimental methods and procedures to insure proper control of the testing procedure. The last two studies—those dealing with sponges and cereal—are like those often conducted by advertising agency research departments and by marketing executives on the staff of many manufacturing concerns. Long practical experience has taught them, or they have learned from competent authority, many of the experimental controls familiar to the psychologist.

There have been many product test studies conducted by naive experimenters whose methods were faulty. But not all of these studies were necessarily misleading in their results, since taste, or color, or other preferences may have been sufficiently large in one direction to have been proof even against faulty methods. But in those instances where results were inconclusive or misinterpreted, the waste in time, effort, and dollars due to faulty method has probably been great.

Practicality is essential for the psychologist in his conduct of a product testing study or in counsel in a product test situation. The psychologist's milieu is frequently the laboratory, but that will not necessarily be the proper place to conduct a product test study. Many complications could arise in measuring taste preferences for oatmeals in the laboratory, but a measure of the extent of discrimination of differences in preparation would definitely belong there. Frequently the number of subjects is important in product testing because many variables, age, sex, previous conditioning, etc., affect consumer preferences. Psychologists studying the fundamentals of human behavior in their own laboratories frequently use many trials with a few subjects. When studying preferences for products, many subjects with a limited number of trials is frequently the best procedure. Such studies are often more economical and more fruitful outside of a laboratory. Many dollars are frequently at stake in product testing. Before it is undertaken, it must be carefully planned so that it will produce results of sufficient value to justify the expenditure necessary to conduct it.

READING REFERENCES

Ripin, R., and P. F. Lazarfeld, "The Tactile-Kinaesthetic Perception of Fabrics with Emphasis on Their Relative Pleasantness." *Journal of Applied Psychology,* 1937, **21,** 198–234.

Schlosberg, H., "A Comparison of Five Standard Shaving Creams by the Method of Constant Stimuli." *Journal of Applied Psychology,* 1941, **25,** 401–407.

54. COPY TESTING

ALBERT D. FREIBERG, PH.D.

Director of Marketing and Social Research Division, Vice President, and Member of Board of Directors, Psychological Corporation; Fellow, American Psychological Association; Lecturer, Columbia University; Diplomate in Industrial Psychology of the American Board of Examiners in Professional Psychology.

Copy testing is an attempt to give an accurate answer to questions such as: How many people will see and read my advertisement? How many will listen to my commercial? How many will understand and remember one or more ideas? How many will buy my product or think favorably of my company?

During the past twenty years the field of copy testing has become increasingly active and important. Nevertheless, there is still considerable lack of knowledge, and disagreement about the various techniques and the uses to which they may be put. This chapter will discuss briefly the principal copy-testing methods used today in testing magazine and newspaper advertisements.

READERSHIP TEST (RECOGNITION)

This test consists essentially of going through a magazine or newspaper with a reported reader and recording what the respondent says he has looked at or read. Usually several hundred or more reported readers are interviewed concerning one issue. Results are then recorded in terms of the percentage of readers who have looked at or seen an advertisement, or such parts of it as headlines, illustrations, logotype; who have read part of the advertisement; and who have read most of the advertisement.

RELIABILITY AND VALIDITY

Split-half correlations of reliability of the readership test range around + 0.90, which indicates adequate stability of responses. Validity is more difficult to determine. Respondents often say that they "recognize" materials which they have not seen before. This has been demonstrated many times by inserting an advertisement that has not appeared in a previous issue of a magazine.

D. B. Lucas has probably done more than anyone else to point out the exact nature of these errors, their importance, and the factors which are responsible for them. He has also devised the Controlled Recognition Test (Lucas, 1940), the essential characteristic of which is the testing of advertisements or editorial materials, before they have appeared, and later after they have appeared. This makes possible a correction of the postpublication test scores for each item tested in terms of the prepublication test score for that particular item.

The formula devised by Lucas (1940) is:

$$\%\ \text{Readership} = \frac{\%\ \text{YES on posttest} - \%\ \text{YES on pretest}}{100 - \%\ \text{YES on pretest}} \times 100.$$

For example, if 20 per cent of the respondents said they saw a particular advertisement by Ford before the advertisement had been published (pretest) and 40 per cent said they had seen it when it was later tested, after publication, the score, instead of being 40 per cent (posttest score) would be:

$$\frac{40 - 20}{100 - 20} \times 100, \text{ or 25 per cent.}$$

A study reported by Lucas and Murphy (1939) is instructive in this connection. They found striking differences in the percentage of respondents who said they had seen an advertisement although it had not yet appeared in a publication. This incorrect recognition occurred with all kinds of advertisements—small and large, of different products, in the front and the back of the book, and the like, as is shown by the following data taken from their study:

Size of Advertisement	*Average Incorrect Recognition before Publication* (Percentage)
Two pages (20 ads)	20
Full pages in color (60 ads)	24
Full pages in black and white (94 ads)	20
One-half page (53 ads)	20
One-quarter page (50 ads)	17

More significant than these averages is the range of incorrect recognition, which varied from 0 for a black and white Parke, Davis full-page advertisement to 58 per cent for a colored Coca-Cola full-page advertisement. The following data, also from Lucas and Murphy, shows the range of incorrect recognition.

Range of Prepublication Recognition Scores (Incorrect Recognition)

Percentage of Readers Recognizing Ads	Number of Ads
10 or less	44
10 to 20	96
20 to 30	85
30 to 40	30
40 to 50	16
50 or more	9

These results agree with the findings of other experimenters with regard to the great variation in the amount of incorrect recognition of different advertisements. It is immediately obvious that any correction on the basis of some constant, for example, 5 or 20 per cent, or by a proportion of the postpublication score, for example, 20 per cent of the postpublication score of 30 per cent, is inadequate. Probably the best corrective factor would be based on the prepublication recognition score of a particular test item by means of a formula such as Lucas proposed. Unless scores can be corrected in some such manner, an advertiser who gets a recognition score of 50 per cent for his advertisement cannot tell whether the error is 10, 20, or 30 per cent.

Interesting in this connection is the following table from Lucas.

Correlations between Ordinary Recognition Scores and Controlled Recognition Scores for the Same Magazine Advertisements

Correlation Coefficient (*Pearson r*)
Recognition Scores
— by Leading Commercial Recognition Service
— and by Controlled Recognition Method

Saturday Evening Post (1939)	Number of Test Exhibits	*Commercial Ratings Compared with*			
		Uncorrected Recognition		Corrected (Controlled) Scores	
		r	*P.E.*	*r*	*P.E.*
March 18	23	+.85	±.04	+.66	±.08
March 25	20	+.80	±.05	+.56	±.10
April 1	20	+.75	±.07	+.61	±.09
Combined three issues	23	Full pages in color		+.38	±.12

There may be a spurious element in correlating a series of advertisements which range from black half-pages up to the colored full pages. When the correlation is based only on colored pages in the three issues, it drops to + .38, as shown above.

Part of this erroneous recognition is due to deliberate falsification on the part of the respondent, who feels that he should have seen or read a particular advertisement or editorial feature. Sometimes the erroneous recognition is due to the similarity of the item tested with one the respondent has actually seen. There is considerable similarity between advertisements, particularly in a campaign which extends over a period of time. Prepublication scores of advertisements in a campaign of similar advertisements show an increase as the campaign progresses.

Recognition and identification scores are most valid and reliable for newspaper editorial items since the testing is usually done within 24 hours after the reading of the material. Test scores for magazine editorial items also are valid because a story or article has unique identifying characteristics. Test scores for newspaper advertisements are somewhat less valid; and least valid of all are test scores for magazine advertisements.

Despite, however, the shortcomings of the recognition technique in measuring the readership of a particular advertisement, it has been and is successfully used by advertising men in the evaluation of advertising. A knowledge of error allows the results of tests to be interpreted accordingly.

Among the questions which may be answered by means of readership tests are: Which magazine or newspaper will give me the most for my dollar? What is the best type of advertising for my product, comic or display? Should I use a large or a small space? Is my campaign gaining or losing interest? Are people more interested in the campaign of a competing brand?

RECALL TEST

Recall tests assume various forms, but in contrast to the Recognition Test, the advertisements or test materials are not shown.

UNAIDED RECALL

Early studies of the recall value of advertising (Gale, 1900; Scott, 1908) used unaided recall tests, *e.g.*, a person who said he had read an issue of *The Saturday Evening Post* was asked, "What advertisements did you see?" Results from this test are too general to be of use in solving most advertising problems.

AIDED RECALL

This test may be illustrated by the question, What brand of coffee have you seen or heard advertised lately? It is immediately clear that this test is intended to measure comparative brand dominance. It indicates the brand name uppermost in the minds of the men and women interviewed. Since factors, such as the use of and prestige of a particular brand influence results, the relative dominance of various brands, as indicated by this test, does not necessarily correlate highly with the relative sale of brands. However, this test indicates trends in the effectiveness with which advertisers are making an impression on the minds of their prospective purchasers.

In this connection the results of two studies made by The Psychological Corporation 12 months apart are interesting (Freiberg, 1946). Each study was made in more than seventy-five cities and towns in the United States, distributed properly by city size, geographic area, and socio-economic groups. Half the interviews were made with men and half with women. The question asked was, "What brand of dentifrice have you seen or heard advertised lately?" Results follow.

Brand	*October 1937*	*October 1938*
Pepsodent	39.4%	25.7%
Ipana	18.4	18.6
Dr. Lyon's	13.7	17.4
Colgate	6.2	16.7
Listerine	4.2	4.9
Kolynos	3.5	4.1
Etc.		
Total	100.0%	100.0%
Total Number of Interviews	5,000	5,000

It is clear that in the 12-month period between these two studies Pepsodent dropped from 39.4 per cent to 25.7 per cent. Most of this loss was due to a decrease in the effectiveness of radio advertising. Another significant change is the gain made by Colgate from 6.2 per cent to 16.7 per cent. A special analysis showed that this gain was due to Colgate radio advertising during this period. This table shows strikingly that radical changes in brand dominance are possible in a period of 12 months as a result of powerful advertising influences. Certainly the advertising managers of Pepsodent and Colgate would have different reactions to these results. Up-to-date data of this kind are useful for evaluating past advertising and also in making plans for future advertising.

After he had named the brand of dentifrice he had seen or heard advertised, each respondent was also asked, "What did this advertising say?" The following data from the same study of The Psychological Corporation shows the responses given by the men and women who named Pepsodent:

Features of Pepsodent Advertising Recalled	*Oct. 1937*	*Oct. 1938*
Brush your teeth twice a day (and/or) see your dentist twice a year	19.7%	10.1%
Contains Irium, new chemical, new ingredient	35.8	26.0
Gives you whiter teeth; keeps teeth white	10.9	11.6
Removes film; removes tartar	6.0	4.1
Keeps teeth clean, makes teeth brighter, makes them glisten, preserves teeth, good for teeth	6.9	16.0
Amos'n'Andy; radio program (no specific feature)	5.5	7.0
Etc.		
Total	100.0%	100.0%

There are several points of interest here. First of all, by October 1937, Pepsodent had succeeded in impressing the idea of using Irium, a new chemical ingredient. The second idea which was mentioned by a large proportion of the respondents was the idea, "Brush your teeth twice a day; see your dentist twice a year." This theme had originally been plugged heavily by Pepsodent but had not been stressed for some time preceding October 1937. The result, which showed that the idea maintained itself without the aid of active promotional effort, indicated that the idea itself was a good one. By 1938, a year later, the popularity of this theme had dropped to 10.1 per cent. Even a well established association needs reinforcement in order to maintain itself forever.

In response to the question, "What did this advertising say?", those individuals who said they had seen or heard Dr. Lyon's advertised gave the following responses:

Features of Dr. Lyon's Advertising Recalled	*Oct. 1937*	*Oct. 1938*
Clean your teeth with powder; do as your dentist does	69.6%	75.4%
Gives whiter teeth, keeps teeth white	8.0	4.5
Keeps teeth clean, makes teeth brighter, makes teeth sparkle	5.6	4.2
Use powder (instead of paste); powder is best	7.3	5.5
Etc.		
Total	100.0%	100.0%

"Clean your teeth with powder; do as your dentist does" is the essential association established by Dr. Lyon's. Independent tests made by The Psychological Corporation between 1932 and 1938 had shown a steady increase in the number of people who associated this idea with Dr. Lyon's. The sales records of Pepsodent during this period were downward; Dr. Lyon's showed a steady increase in customers.

TRIPLE ASSOCIATES

Much advertising is developed around a central theme or idea. Here a special form of recall test developed by Link (1933) may be used to good advantage. It may be illustrated by the test question, "What automobile company advertises: 'First by far with a postwar car'?" A nation-wide test by The Psychological Corporation based on 2,500 interviews with urban men and women properly distributed by geographic area, size of city, and socio-economic groups gives the following results as of May 1947, approximately a year after Studebaker started using this theme:

Studebaker	16%
Ford	11
Kaiser-Frazer	3
Others	8
Don't know	62
Total	100%
Total Number of Interviews	2,500

In situations of this kind this test is excellent for measuring the extent to which an advertiser has succeeded in associating his name with a particular idea or theme.

Some years ago The Psychological Corporation was asked to test a newspaper campaign of a dishwashing soap. The principal components of the plan of advertising were: A company representative made arrangements in a town for one of the local home economics instructors to give a lecture on cleanliness in the home to a group of local women, usually a church or club group. At the termination of the lecture the company representative gave a demonstration of the effectiveness of his brand of soap, and compared it with several other leading brands. The test consisted of putting a tablespoonful of each brand of soap in each of several quart jars and then adding a cup of boiling water. The net effect was a certain amount of soap suds and, interestingly enough, the advertiser's soap produced more and deeper suds than did any of the other leading brands. Local photographers then took pictures of the demonstration, including in the picture one or more of the prominent local club or church women. This picture was then incorporated into the local newspaper advertisement of the soap using the headline, "The Deeper Suds Test."

After looking at the advertising, a Controlled Opinion Test was made comparing the Deeper Suds Test advertisement with that of other leading brands of soap. The results indicated that the Deeper Suds Test advertising interested housewives only moderately. More significant was the finding that upon blocking out the name of the advertiser there was considerable confusion with Super Suds, a competitive brand. The obvious recommendation was made that the theme, "The Deeper Suds Test," be dropped. However, the client wanted further evidence, and accordingly, the following Triple Associates Tests were carried out:

An 8-week newspaper campaign was scheduled in each of four cities. Preceding the first advertisement, interviews were made with 300 housewives of the middle socio-economic group in each of these cities, a sampling of the kind of women who buy soap for dishwashing. Interviews were widely scattered in each city and women in even-numbered homes and apartments were interviewed on the pretest. Two questions were asked: (1) "What brand of soap do you use for dishwashing?" and (2) the Triple Associates question, "What brand of dishwashing soap advertises *The Deeper Suds Test?*" The results were:

Brand Named	*Test Made before Advertising Campaign*
Super Suds (incorrect)	12%
X—the advertiser!	1
All other brands	3
Don't know	84
Total	100%
Total number of interviews	1,200

The client was again warned that the greater share of his advertising efforts was developing a market for Super Suds, a competing brand, but he refused to stop the advertising campaign. As can be seen below, the eight weeks of newspaper advertising did help Super Suds. In the second, or post-advertising test, interviews were made with women in odd-numbered homes and apartments, directly across the street from or alongside those interviewed in the first study.

Brand Named	*Test Made after Advertising Campaign*
Super Suds (incorrect)	16%
X—the advertiser!	2
All other brands	3
Don't know	79
Total	100%
Total number of interviews	1,200

This study shows several things about the Triple Associates technique. First, it can be used prior to any advertising in order to discover the amount of confusion, if any, which exists between the central idea or theme the advertiser proposes to use and the advertising of other brands. There is usually only a small amount of confusion concerning most advertising themes or ideas. A second observation that can be made from this study is that the large proportion of people who do not associate the theme or idea correctly with the advertiser say they do not know, that is, there is usually only a small amount of incorrect guessing.

RELIABILITY AND VALIDITY OF RECALL TESTS

These tests show split-half reliability coefficients of approximately +.90, comparable in size to those of the readership tests. If the purpose of recall tests is the measurement of comparative brand dominance, as well as the measurement of the association of ideas with the brand name, then the validity of recall tests is self-evident. It must be remembered, however, that this association is not entirely due to advertising only, but also to the use and familiarity with the product over a period of time. Because of this it is often difficult, if not impossible, to isolate completely the effects of advertising as such.

USES OR APPLICATIONS

A great deal of advertising proposes to establish an association between a need or a want and a particular brand or make of product, *e.g.,* "We need a new automobile—let's buy a Studebaker." "I need a new hat—I'll buy a Stetson." The advertiser who can establish this kind of association will have predisposed the potential customer toward the purchase of his product. Other factors, such as availability, special sales, and price, also influence purchase of a product. These tend to reduce the correlation between actual sales and the results of recall tests.

The practical value of the recall test depends on the care with which the results are interpreted. Recall tests are of greatest value when repeated periodically so that an advertiser can determine his comparative ranking in relation to his principal competitors, as well as any increase or loss of his own standing. Periodic studies indicate the extent to which his copy themes or ideas and those of his competitors are making headway or losing ground.

OPINION TESTS

Opinion testing consists essentially in showing one or more advertisements to an individual and then asking certain pertinent questions about the advertisements. If more than one advertisement is shown, an individual may be asked, "Which of these advertisements interests you the most?" Or, "Which illustration most appeals to you?" Or, "Which headline do you like best, and why?" Sometimes the advertisements are taken away and the respondent is asked, "What is the principal idea the advertiser is trying to get across?" This test is a direct and immediate measure of people's interests, likes, and dislikes.

RELIABILITY

Reliability of opinion tests is satisfactory. When testing imitative and actual advertisements, Strong (1911) secured repeat reliability coefficients ranging from 0.84 to 0.95 with various groups of subjects as small as 25 in number. Newhall (1930) reported 64 repeat reliability coefficients with groups of from 12 to 41 subjects ranging approximately from 0.50 to 1.00. All correlations were positive, 91 per cent were above 0.50, 80 per cent above 0.60, 66 per cent above 0.70, 55 per cent above 0.80, and 23 per cent above 0.90.

VALIDITY

Starch (1923) reported positive correlations between test rankings by this method and sales and inquiry returns ranging from 0.40 to 1.00, with an average of 0.79 for twenty series of advertisements. Borden and Lovekin (1935) reported a study made on seventeen series of magazine advertisements in which each series contained from two to five advertisements which showed reliable differences in sales or number of inquiries. The men and women tested were from economic, social, and educational groups which represented consumers for the advertised products. When the article advertised was limited in appeal, rankings between inquiries or sales and opinions agreed in only 36 per cent of the cases, but when the advertised article had wide appeal, opinion and inquiry or sales rankings agreed 92 per cent of the time. These results indicate that a representative consumer group must be tested for results of maximal validity.

USES AND APPLICATION

One of the advantages of this technique is that it can be used to test advertisements previous to their publication, before large amounts of money have been spent. In testing, finished art work is not necessary and rough layouts can be used with typed copy stripped in. This makes pretesting comparatively inexpensive. In addition to the main purposes of the testing, many incidental results of importance come out. For example, a test of three anti-freeze campaigns indicated that less than 5 per cent of the men car-owners tested mentioned the scientific name for one of the ingredients in the anti-freeze despite the fact that this scientific name had been given prominent space in the advertising treatment in one of the campaigns. This result, incidental to the main purpose of the test, led the advertiser to eliminate this copy and insert other hard-selling copy in the space.

Some years ago a leading canning company advertised full-page in Sunday newspapers a contest featuring ten first prizes of paid vacation trips to Hollywood for the winners. After spending considerable money for this full-page advertising on two successive Sundays, the advertiser became concerned because of the small number of entrants in the contest. As is usual in situations of this kind, conferences were held. It was finally decided that some inquiry should be made to discover why there were so few entrants, and The Psychological Corporation was called in for consultation. A controlled opinion test made locally with the advertisements indicated that almost half of the women who used this particular kind of canned food regularly had small children and therefore would have found it inconvenient to go to Hollywood. Almost another one fourth of the women interviewed did not want to go to Hollywood without their husbands. Another test made at the same time with single women under twenty years of age, showed that more than three fourths of them were intensely interested in going to Hollywood but that only 1 out of 6 had ever prepared a can of the food. This particular advertiser could easily have saved well over $75,000 by a relatively simple opinion test of the basic appeal—the unwanted trip to Hollywood!

A WORD OF CAUTION

This test may be invalid if the advertisements contain claims or themes which are not socially acceptable; for example, if they emphasize sex or morals. Advertisements containing pictures of babies are often overrated because of the high interest in babies. Like all tests, the results are of value in proportion to the cooperation and interest displayed by the respondent. Motivation must be adequate; otherwise the results are of little significance.

INQUIRY TESTS

Many advertising men feel that inquiry tests are particularly valid because the respondent clips a coupon. The overtness of the response seems to imply validity, despite studies which have shown that many respondents are not typical of the population to which the advertiser is appealing. The type of offer is extremely varied. Here are a few of the more common offers:

1. A sample of the product
2. A brochure or booklet about the product; how it can be purchased, etc.
3. A premium or novelty of some kind
4. A special bargain, such as two for the price of one

Sometimes these offers are free, sometimes they require a purchase or an order, occasionally there is a small postage charge. They may even involve sale at full price if it is mail-order advertising.

The offers are also presented in various forms, the most common of which are:

1. A coupon to be torn or cut out, a part of the advertisement itself
2. A buried offer, that is, an offer made in the text, usually toward the end of the copy
3. An invitation for a written request, telephone call, or request to take the coupon to a dealer

VARIABLES AFFECTING INQUIRY REPLIES

It is generally agreed that the number of inquiries is affected by factors such as:

1. *The offer made.* The nature as well as the appearance of the offer affects the number of replies. If the offer is made in a coupon it should always be in the same part of the advertisement and identical in other respects, that is, in appearance and wording. If it is a buried offer then the wording and relative position in the copy should be the same.

2. *Page position.* The advertisements should occupy the same space since page position and position in the magazine or paper affect pulling power.

3. *Size of the advertisement.* The advertisements should be of the same size.

4. *Media used.* The media used should reach the same kind of people which the advertiser intends to reach in the final campaign.

5. *Time interval between advertisements.* This interval must be long enough so that the majority of replies can be attributed to a particular advertisement.

6. *Competition from news.* Tests often suffer from competition with unusual news, such as elections.

7. *Seasonal factors.* A series of tests may take several months to complete. During this period seasonal variations in the weather may affect the results.

8. *Competitive factors.* During the test period there may be decided changes in competitive merchandising efforts. Competitors might even introduce special sales in the test cities in order to "facilitate" the advertising test of another company.

9. *Weather.* The effect of extreme weather may be important in the testing of particular products, such as umbrellas and anti-freeze, whose sales are affected by changes in weather conditions.

In addition to these general factors, there are specific factors which are important in the testing of individual products. These must be considered in evaluating results of inquiry tests.

TESTING BY THE SPLIT-RUN METHOD

Many of the variables just mentioned can be controlled by the split-run method, the runnning of two or more advertisements at the same time in the same part of a magazine's or a newspaper's circulation. For example, advertisement A appears in one third of the newspaper's circulation, advertisement B in another third, and advertisement C still another third. The three advertisements appear on the same page of an issue, compete with the same news and advertising, and therefore are run under approximately equal conditions. In this way many variables, such as weather conditions and seasonal and marketing factors, are kept constant. Therefore differences in number of returns are due primarily to differences in the advertisements themselves. There are now a number of magazines and newspapers which offer this type of advertising.

USES AND APPLICATION

Inquiry tests are used to measure the comparative effectiveness of two or more advertisements, two or more campaigns, etc., in preparation for the running of advertising in more expensive space. Tests are usually made in magazines or newspapers of limited circulation, although recently a few magazines and newspapers with large circulations have offered split-run facilities.

RELIABILITY AND VALIDITY

Given an adequate number of replies, reliability coefficients are usually adequate. More important and difficult to demonstrate is their validity. Even the split-run technique does not prove the validity of the inquiry returns, and correlations of inquiry returns with sales are found to vary considerably.

The closer an inquiry return approximates an actual sale, the more likely it is that the inquiry return gives a valid indication with regard to the variable tested. When inquiry returns are followed up by mail or personal sales efforts the validity of a test comparing two appeals can be demonstrated. Seasoned advertising men have determined, by trial-and-error, the type of offer, media, and other factors necessary to make inquiry returns valid tests for the sales of a particular product.

Most advertising situations do not, however, allow this type of validation, particularly institutional advertising. The adequacy of the inquiry test depends on the personal skill and judgment of the individual making the test. The different abilities of those who do inquiry testing account, in part, for the varying validity of inquiry tests, which may be one reason for the divided opinion with regard to the usefulness of inquiry tests.

SALES TESTS

Since the avowed purpose of most advertising is to produce sales, either immediate or delayed, either directly or indirectly, it is easy to understand the high esteem in which the sales test is held. At first glance a sales test seems simple. Actually it is very difficult to administer because of the many variables other than the advertising which affect sales. Sales tests are very expensive, too. Nevertheless, they are often used.

TESTING BY MEANS OF SPECIAL CITIES

A common variety of sales tests measures the comparative sales effect of two or more campaigns. Usually a limited number of cities is selected for testing, two or three for each of the campaigns to be tested. Control cities are often used; in these either no advertising is run or the current advertising is continued. Cities are selected on the basis of comparability in size, average income of families, comparability of newspaper circulation, types of population, and other factors which have an effect on the sales of the product tested.

Other influences which must be taken into consideration are: weather, epidemics, seasonal changes, the length of the test period, and competition of other advertisers (Coutant, 1932). Unless these variables are either controlled or their effect taken into consideration, sales tests may not be valid. Sales are usually measured before a campaign, while it is running, and at its completion.

METHODS OF MEASURING SALES

Sales are commonly measured in the following ways:

1. *Sales of companies.* Sales records of the company to wholesalers, jobbers, and other clients are used. This is only an approximate measure because of the unknown lag between sales by the company and sales to consumers who are influenced by the advertising. Dealer inventories, warehouse inventories, and uncertainty of area serviced by specific jobbers all contribute to this unknown lag. Usually a more accurate measuring device is used.

2. *Sales at retail outlets.* Selected retail outlets in the test and control cities are inventoried at regular intervals. The A. C. Neilsen Company of Chicago has made an important contribution here in the drug, grocery, and liquor fields. Sales are measured by bimonthly inventories in retail outlets scattered throughout the United States. (The refusal of certain chain outlets to cooperate produces some bias in the typicality of the sample for certain products.) Data are reported by brand, by number of units sold, and also by dollar volume. Many companies use this excellent service for measuring the results of sales tests.

3. *Consumer purchases.* Sales tests are often checked by consumer panels or by periodic studies made in comparable samples. In consumer panel studies members keep records of all purchases. These are collected and tabulated at regular intervals. In periodic studies questions on recent purchase or usage are asked, and sales trends analyzed accordingly. This type of measurement occurs at the terminal point of the advertising, that is, with the consumer. By these methods, therefore, the effect of advertising can be measured in terms of such factors as age, sex, and educational level—all chacteristics important to the advertiser, and thus give information not obtained by other modes of measurement.

WHAT TESTS SHOULD BE USED?

From the preceding discussion it is evident that different tests measure somewhat different things. Therefore, an advertisement or an advertising campaign might not do as well on one test as on another. A table which illustrates this point shows the rank order of eight advertising themes when tested by three aided recall tests, an opinion test, and a recognition test (Freiberg, 1946).

Rank Orders

Theme	Recall Test I	Recall Test II	Recall Test III	Opinion Test IV	Recognition Test V
A	1	3	1	1	5
B	3	1	5	2	6
C	5	2	2	3	2
D	2	4	4	5	1
E	6	5	3	4	4
F	4	6	7	8	3
G	7	8	8	6	7
H	8	7	6	7	8

The recall tests and the opinion test ranked the themes in similar orders. However, the recognition test gave quite a different ranking. Correlations between the results of these tests, computed by the rank difference method, are:

	Recall Tests I	Recall Tests II	Recall Tests III	Opinion Test IV	Recognition Test V
I		0.59	0.40	0.43	0.40
II			0.60	0.76	0.20
III				0.70	0.24
IV					0.21

The ranking in tests I, II, and III are based on 200 interviews; those in the opinion and recognition tests, on 300 interviews each. Since there are only eight ranks in each test, the probable errors of the correlations are high. Nevertheless, the consistent trend shows low correlations between the recognition tests and the recall and opinion tests. In contrast, the correlations between the recall and opinion tests are fairly high, which indicates that the first four tests are measuring factors in common, but not what the recognition test is measuring.

There are, of course, many instances where different tests correlate highly, although this is not the usual result. Since the tests measure different things and the advertisements are not uniformly high or low in all their characteristics, such as attention value, interest value, and action power, one would expect this.

The test used must be selected with relation to the purpose of the advertising. For example, if the purpose of an advertisement is to develop an awareness of a trade mark or a picture of a new package design, then a readership or recall test is indicated. If, however, the purpose of an advertisement is to convince the public that a particular make of automobile has the latest postwar engineering developments, then an opinion test might be the best. The purposes of advertisements and advertising campaigns are not always as easy to classify as this and consequently the selection, use, and interpretation of tests require considerable judgment and skill.

SUMMARY

The tests for advertising copy described vary widely. There are, in addition, other techniques. No one test is the final answer to the problem of the advertising man.

Copy testing, like interviewing and job analysis, is one of the arts of psychological practice. It almost but never completely achieves scientific status. However, test results often make sense, not only to practical advertising men but also to scientifically trained research men. Even in its present ill-defined state, copy testing gives valuable clues to advertising men, partly because many of the problems arising are of even a grosser nature than the copy-testing tools employed to solve them.

READING REFERENCES

The Advertising Research Foundation, *Copy Testing; A Study.* New York: The Ronald Press Company, 1939.

Caples, John, *Advertising for Immediate Sales.* New York: Harper & Brothers, 1936.

———, *Tested Advertising Methods.* New York: Harper & Brothers, 1932.

Firth, L. E., *Testing Advertisements.* New York: McGraw-Hill Book Company, Inc., 1934.

Rudolph, H. J., *Four Million Inquiries from Magazine Advertising.* New York: Columbia University Press, 1936.

55. ADVERTISING MEDIA RESEARCH

DARRELL BLAINE LUCAS, PH.D.

Professor of Marketing and Vice Chairman of the Marketing Department, New York University; Technical Director, Advertising Research Foundation; Research Counsel, Batten, Barton, Durstine & Osborn; Member of the Joint Committee of the National Research Council and Social Science Research Council for the Measurement of Opinion, Attitudes, and Consumer Wants; Member of the Original Life Committee on the Continuing Study of Magazine Audiences.

An advertising medium is the carrier of an advertising message, whose primary functions are entertainment and the dissemination of information. Major media include newspapers, magazines, and the radio. Outdoor displays, transportation advertising, and direct-by-mail advertising are usually classed as media, although they are actually forms of advertising having no other function. Zeisel, in *Printers' Ink,* reported the following estimates of the total investment in each type of advertising for 1947 in the United States as follows: newspapers, $1,222,000,000; radio, $530,600,000; magazines (including farm journals), $492,900,000; direct-by-mail, $482,700,000; outdoor, $121,400,000; business papers, $196,000,000. Miscellaneous media, including transportation advertising and operating costs, brought the total estimated expenditure for the year close to $3,900,000,000.

The functions of an advertising medium are to provide an advertiser with space or time, and to supply him with a potential audience. A medium also may influence the amount of exposure of an advertisement and the favorableness with which it is received. Every medium provides its own peculiar psychological setting for the delivery of advertising messages.

MEDIA RESEARCH

Advertising media research measures the size of the advertising audience, subdivides the audience into various classifications, and appraises the attitude of the audience toward the medium and its commercial messages. More specifically it measures: (1) the total number of individuals receiving a psychological impression from the medium and its advertisements; (2) the number of repetitions and duration of contact; (3) the divisions of the audience into various market groups; (4) audience attitudes toward the medium and toward the advertisements carried by it; and (5) the total advertising effect as related to each specific medium.

TECHNIQUES OF MEDIA RESEARCH

The ultimate criterion of the value of an advertising medium is, in general, its usefulness in aiding the sale of goods and services. The obstacles to the measurement of the selling impact of major media have been so great that most media research is directed toward the intermediate psychological steps that are essential to successful advertising results. An effective medium reaches large numbers of the right kind of people when they are in a frame of mind favorable to the purposes of the advertiser. A good medium presents each advertisement in such a way as to encourage a maximum number of people to see it or hear it and then to read it or listen to it. This obvious emphasis on psychological effects has brought psychological research techniques into the forefront of media research.

Outstanding among research techniques is the *recognition method* developed by Gallup on newspapers and first applied to magazines by him in 1931. This is called a recognition technique because it depends upon the recognition of items previously seen or read. It is not primarily a measure of memory as such, but rather a method of determining how many people have previously seen or read each item. In essence, typical readers of a publication are interviewed several days or weeks after an issue is published, and are asked to go through the issue pointing out what they have seen or read.

While the recognition method has achieved greater usefulness than any other technique applied to advertising and media, there are several other techniques of genuine value, most of them being largely of a psychological nature. The measure of an advertising audience is usually applied coincidentally—at the time the advertising is seen or heard—or at a later time when some form of recall is required. The measurement of attitudes poses a far more difficult problem and rests almost exclusively on interviewing techniques evolved through psychology.

SIZE OF AUDIENCE

MEDIA AUDIENCES AND ADVERTISING AUDIENCES

The comparison of advertising media, based upon their audiences, involves both the audiences of the media and the audiences of the specific advertisements which they carry. The audience of a specific advertisement is necessarily a part of the total audience of the medium. In the case of a magazine, for example, a certain number of people open an issue and look through a part of it. These people are the audience of that issue of the magazine, but it is unlikely that all of them will turn the page upon which a particular advertisement happens to appear. Therefore, the advertiser has no chance to reach all of the people in the magazine audience. On the other hand, the advertising message may be of such a nature that many people who turn the page give it scant attention or none at all. Thus the actual audience gained by the advertiser is partly his responsibility as well as that of the medium. Any attempt to evaluate media and to compare them on the basis of audiences must take account of the part played by the medium and by the advertisement in gaining the audience of a specific advertisement.

The problem of evaluating the radio in terms of advertising audiences is unique. The greater share of all listening is devoted to programs sponsored by individual advertisers for business purposes. The audience of a specific program may belong to the sponsor, but it is clear that some of the listeners may not hear any of the commercial messages. On the other hand, mere awareness of the sponsor's identity may reflect upon him a certain amount of good will. It is more difficult to define just what is the total audience of the radio as a medium and the audience of radio advertising than is the case with publications.

RELATIVE RATINGS OF PROGRAMS AND ADVERTISEMENTS

One of the earliest attempts to make a systematic appraisal of the audience of a medium was the establishment in 1930 of the Cooperative Analysis of Broadcasting. This organization obtained percentage ratings for the audiences of individual radio programs by means of telephone interviews in certain large cities. Its chief technique was known as the day-part method in which people were asked to report, without memory aids, the programs heard within two hours previous to the call. C. E. Hooper, Inc., in 1934, established a rating service, limited to telephone homes in large cities, using a technique called the coincidental method. Instead of relying upon memory, the coincidental interview concerns itself only with what people are listening to at the moment of the call. Both of these procedures arrived at scores based upon such a limited segment of the population as to be of much more value for relative comparisons than for estimating the total number of actual listeners. Therefore, the listeners per dollar could not be computed.

Likewise, Daniel Starch, in 1932, initiated a rating service for magazine advertisements based upon an application of the recognition method to a limited selection of magazine readers. Readers of a magazine are selected, partly upon the basis of convenience, and are asked to point out the advertisements which they have seen and to indicate the degree of their reading. The ratings show the percentages of the readers of the magazine who see and the percentages who read each of the advertisements. Despite the value of these ratings in comparing various advertising techniques and despite the inclusion of figures on cost-per-hundred readers, these ratings, like the radio ratings, are generally considered inadequate for projecting total advertising or media audiences. This is not said to disparage the value of any of of these types of ratings, but it indicates why so much later effort has been expended in order to obtain more acceptable audiences of media and of advertisements.

MEASURING THE TOTAL AUDIENCE OF A MEDIUM

The total audience of a national advertising medium, such as a magazine, a radio network, or a possible newspaper combination, may include almost anyone within the family units which are reached or, indeed, in the population of the country. Usually advertising media audiences are limited arbitrarily to persons above a certain minimum age. Such media as newspapers, radio stations, and transportation advertisements are limited in their coverage to smaller geographical areas such as a single city. The audience of the medium and, therefore, the potential audience of the individual advertisers, becomes a large share of the total population of the area.

So long as the area covered by a particular advertising medium coincides with areas of known census figures, the audience of the medium may be measured by determining the percentage of the known population which is reached. On the other hand, if the medium reaches but a small percentage of the total population of the area served, it may be simpler to measure the audience in terms of the families or individuals served in the area.

Life Magazine, in 1938, initiated a national Continuing Study of Magazine Audiences, the purpose of which was to determine the number of people, ten years of age and older throughout the United States, who read or look through a typical issue of each of the leading weekly magazines. As the study was originally conceived, it was desirable to select a representative sample of the total population ten years of age and older, and to determine whether each person had read or looked through a specified magazine issue within approximately four weeks of the date of issuance. Because of the danger of mistaken identity, and its inflationary effects, *Life* adopted a form of controlled recognition to avoid this inflation. Copies of the magazines were taken into the field before their date of issuance to determine what percentage of the population sample would claim to have read the issues to which they had not been previously exposed. The resulting percentages of mistaken identification were used to reduce and adjust the later reading claims made after publication of the magazine issues. The figures in Table 1 are taken from *Life's* No. 7 report, together with privately communicated data, and are based upon a total population of 108,600,000 civilians ten years of age and older.

Table 1: Weekly Magazine Audiences (Civilian), 1943

Magazine	Claimed Audience	Advance Claims	Adjusted Audience	Projected Total Audience in US
Collier's	12.8%	3.9%	9.3%	9,850,000
Liberty	9.4	3.9	5.7	6,200,000
Life	26.7	7.1	21.1	22,050,000
The Saturday Evening Post	14.0	3.6	10.8	11,250,000

The adjusted audience percentages in Table 1 are a little larger than the difference between the preceding two columns. This figure is arrived at by a formula which gives a truer measure of the real audience than is indicated by simply subtracting the mistaken "advance claims" from the later "claimed audience." The final projections represent the total number of persons who look into or look through a typical issue of any one of the four weekly publications. They are the audiences of the magazines within which each advertiser builds up his own smaller audience. Similar data have been obtained through privately sponsored surveys of the women's service magazines and the general monthly magazines.

The Broadcast Measurement Bureau, in 1946, published findings intended to represent total radio audiences throughout the United States, county by county. These audiences are defined as the percentage of radio-owning families in a county which listens to a particular broadcasting station as often as once a week. Data are obtained by mailing survey forms to a preselected sample of families and asking them to indicate stations listened to and the frequency of listening. This technique is based on recall. These audiences, based upon mailed replies, indicate a maximum potential size of audience, although they are not strictly comparable to the magazine audiences just described.

The A. C. Nielsen Company measures radio set operation through the use of audimeters which record the particular time periods during which a radio receiving set is tuned to each broadcasting station. This procedure does not give actual evidence of listening behavior and must be supplemented by further research in order to determine the audiences. At this writing, both A. C. Nielsen and C. E. Hooper are expanding their radio research operations in order to achieve ratings capable of being projected to total nation-wide audiences. Since the Hooper telephone coincidental method is not adaptable to most American homes, it is being supplemented by the diary technique. A radio diary is a record form on which the listener keeps a continuous record of his listening behavior—noting the time and the station whenever he changes the radio tuning.

LOCAL AND SELECTIVE MEDIA AUDIENCES

The audience of any medium which can be measured nationally can, of course, also be measured within any city or other census area. The Pulse uses a method capable of measuring the audiences of radio programs in the cities which are surveyed. Interviewers call personally upon a sample of homes in a city and show the respondent a printed list of radio programs called a roster. This method of aided recall is used, with some refinement, to measure listening in all types of homes and to describe the characteristics of the listening audience.

Experimental studies and undoubtedly private studies have been conducted to determine the total number of readers of a particular newspaper within the city or area served. Other studies have been conducted to measure the audiences of outdoor posters and of transportation advertisements in single cities. Outdoor and transportation advertising audiences are, however, the audiences of specific advertisements and will be discussed again in that context.

It seems less logical to measure the audiences of selective media, such as farm papers, business papers, professional publications, and direct-by-mail advertising through samples directly relating to the census. The Continuing Study of Farm Publications, beginning in 1946, measures the audience of a farm publication within the homes of the families receiving it through subscription. Since the publication is circulated almost entirely to known addresses, and since the reading seemed likely to fall within the subscriber families, the audience measurement was considerably simplified by this approach. General magazines frequently have a large percentage of newsstand sale and cannot be sampled so easily. Also each copy is likely to be read by persons in more than one family. Selective media are much more likely to go to known addresses and to be read by individuals or family groups or company groups than is the case with professional magazines, farm publications, and business papers. The audiences of all three of these types of publications are currently being explored or measured by the Advertising Research Foundation.

DETERMINING WHO IS A READER

The method of deflating magazine reading claims by measuring the issues in advance of publication fails to identify who are the true readers. The report of *Life's* "Continuing Study of Magazine Audiences" (No. 8, 1946) is based upon a procedure which reduces mistaken reading claims through a series of questions. The persons who hold to their original reading claims are assumed to have actually read the publication and are thereby identified as readers. The Continuing Study of Farm Publications attempts to achieve the same goal more objectively and more quickly by means of a qualifying "kit" of editorial items, which consists of ten editorial items selected from the measured issue mixed with ten prepublished editorial items from a forthcoming issue. All of those who claim to have read the measured issue of the farm publication are asked to pick out from the kit those items which they have previously seen or read. "Genuine" readers are expected to identify more published editorial items than those not yet published. This procedure not only provides a method of identifying genuine readers for further interviewing purposes, but it also provides a means of calculating the total number of readers of the publication.

ADVERTISING AUDIENCES

That part of the audience of an advertising medium which is important to the advertiser is the one on which the advertisement itself makes an impression. It was said that the advertisement contributes in large degree to determining how many people and just what kinds of people see it. The medium, in still larger degree, determines how many people may come in contact with the advertisement. Even though the advertisement is partly responsible for the actual size of audience obtained, the advertiser judges the medium according to the size of the audience he reaches. If the advertisment is unable to impress all of the people brought within range of the medium, the advertiser at least wants to know what is the maximum audience he can hope to obtain. Under the present limitations of research, the advertiser has no better measure of the dollars he spends on a particular medium than is the maximum audience delivered by that medium to particular advertisers.

THE CONTINUING STUDY OF TRANSPORTATION ADVERTISING

The first advertising medium to measure total audiences, as related to specific advertisements, was transportation advertising through its Continuing Study of Transportation Advertising begun in 1944. Car cards, or transportation advertisements, are displayed 30 days in the vehicles of a particular transportation system. The total number of persons seeing each advertisement is measured and published as a projected figure which may be compared directly with the actual cost of the advertising space.

If the claims which people make about reading a particular magazine, based upon recognition, are unreliable, it seems even more likely that there is error in claiming to have seen a particular advertisement in a single publication, or place of display. It was for this reason that the Continuing Study of Transportation Advertising incorporated the controlled recognition method. This variation of the recognition method requires a complete survey of the advertisements in advance of their appearance in the vehicles. Again, after being displayed 30 days, the advertisements are removed from the vehicles and another complete recognition survey is made.

Table 2. Projections of Chicago Transportaion Advertisements, Based on Basic Population of 2,930,000 People Fifteen Years of Age and Older

Name of Advertiser and Subject of Advertisement	Per Cent of Riders			Percentage of Basic Population	Number Seeing Ad
	Claimed Seen — after 30 Days	Claimed Seen — in Advance	Corrected Real Audience		
Charles Marchand Company Hair rinse	52%	25%	36%	28%	840,000
Wm. Wrigley, Jr., Company Gum	40	11	32	26	770,000
Cutter-Crossette Company Neckties	55	39	26	21	620,000
Marshall Field & Company Department store	25	9	17	14	420,000
Miles Laboratories, Inc. Vitamin capsules	41	33	13	10	310,000

The preexamination scores are used to correct the postexamination scores on the same advertisements, using the controlled recognition formula. The result is an audience percentage based upon the riders of the vehicles, which can then be converted to a percentage of the total population and finally to a projected total audience. Table 2 is an example showing how the audiences are projected.

EXPERIMENTAL MEASURES OF ADVERTISING AUDIENCES. Both the Traffic Audit Bureau (1946) and Point of Purchase Advertising Institute (1946) have done experimental studies very much like the above, in order to measure the total audiences of advertising displays in single cities and the surrounding areas during a typical display period. These studies have shown the feasibility of applying such measures and producing projectible audiences which can be compared with the known costs of the advertisments. A very similar experimental study was conducted by Alfred Politz Research (1946) in St. Louis, Missouri, in order to measure the total audiences of radio commercial messages broadcast on a regular schedule in that city. The method of measuring the claimed audience before any of the commercials were broadcast, and the later claims after a series of regular broadcasts had been made, produced net audiences which are presumed to be as valid as those obtained on display advertisements.

The Nielsen Audimeter offers another avenue for measuring the audiences of commercial messages over the radio. It has already established strong negative evidence that there is little tendency for radio listeners to change the tuning of their sets at the times when commercial messages are broadcast. It remains to be demonstrated just how many people are actually present while the commercial messages are broadcast, and how effectively these messages are impressed upon them.

REPEATED ADVERTISING IMPRESSIONS

The above discussion has taken no account of the duration of advertising contact or the number of repeated impressions of the same advertisements. If a commercial message is given over the radio only once, it can only be heard once. If it appears in a printed form, the same copy may be seen repeatedly. If the same message appears in two magazines, the same person may see it in both. If it appears in successive issues of the same magazine it may be seen repeatedly by the same persons. If, in addition to magazines, newspapers are introduced, the likelihood of repeated impressions is increased still further. Likewise, the introduction of the radio and other media simply multiplies the opportunity for repeated impressions on the same individuals. Certain media, notably outdoor and transportation advertising, are particularly adapted to the repetition of advertising impressions on the same individuals. An outdoor poster is usually displayed for at least thirty days during which time a person may pass the same poster on many occasions. He may pass the same poster in several different panel locations. The problems of evaluating the overlapping of audiences and the advertising value of repetition add complications to the evaluation of media audiences. Their solution is the object of a considerable amount of current research.

AUDIENCE ATTITUDES

Any attempt to describe the people who make up an audience or to appraise their attitudes is usually called qualitative. The greater share of all media and advertising audience studies provide breakdowns of the audiences according to sex, age, economic level, and other characteristics of marketing significance. The introduction of educational levels may be intended partly to aid the advertiser in the plan of his messages and to guide the editor or the radio script writer.

Audience breakdowns are capable of having great psychological significance, but media investigators have for the most part left these interpretations to psychologists and sociologists. There are a number of psychological analysts in the commercial practice of analyzing publications and radio programs, including their advertising messages, and arriving at psychological appraisals of the contents. Their methods vary all the way from completely objective approaches to predominantly subjective appraisals.

ATTITUDES TOWARD MEDIA AND TOWARD ADVERTISEMENTS

McCall's *Magazine* (1939) made its first *Qualitative Study of Magazines*. This pioneering study, together with repetitions in 1941 and 1946, remains the outstanding example of published research on attitudes toward a particular advertising medium. The questioning of women on magazines developed according to three stages, as follows: (1) each woman was asked to state her four leading interests in magazine reading; (2) the respondents were asked to show the actual copies of current issues of magazines in the home; and (3) they were asked to indicate the intensity of their interest in receiving each new copy of each magazine. One might infer that a woman who expressed a particular reading interest *X*, who possessed magazine *A*, and who looked forward eagerly to receiving magazine *A*, might be reading magazine *A* because it satisfied her reading interest *X*. The investigators cautioned against unwarranted inferences regarding the expressed interests, but these patterns correlated so closely with editorial emphasis in the various publications that both the technique and the implied conclusions appear well founded.

Examples from the third McCall study (1946) show that *Good Housekeeping* stood at the top in reader interest among readers of *Good Housekeeping* who were particularly interested in food. *Better Homes & Gardens* stood at the top among readers of that magazine who were particularly interested in home decoration. *Time*, while not a woman's magazine, stood at the top among women readers of *Time* who were particularly interested in current events. *McCall's*, sponsor of the study, failed to gain an undisputed first place but showed great strength in style, beauty, fiction, food, and home management.

Attitude studies toward other media have not been neglected, but have generally failed to evolve techniques capable of wide application or great refinement. The radio, because it invades the public domain, has continuously been the object of attitude studies, particularly of the negative variety. This statement applies to commercial messages as well as to the programs themselves. Outdoor advertising has faced a similar problem though on a smaller scale. Certain magazines have faced opposition in some states, and have met with postal restraint. On the more positive side, media have conducted attitude studies for competitive purposes. When a newspaper strike, 1946, halted delivery around metropolitan New York, a survey was made to learn whether people depended more upon the newspapers or upon the radio for news. Little was contributed to media research techniques.

ATTITUDES TOWARD ADVERTISEMENTS

While a large part of the testing of advertisements deals with reader and listener attitudes, comparatively little research has been done to determine the effect of the medium upon the attitude toward its advertising messages. The McCall study, 1946, showed interest in advertisements to rank in seventh place among the ten leading interests which prompt women to read magazines. *Better Homes & Gardens* was the magazine which, among its own readers, had the highest interest level on the part of those who listed advertising as a major interest. Here again, the inference is that *Better Homes & Gardens* either carries a type of advertising or provides its advertisements with a frame of reference which appeals strongly to the interests of women readers. Many attempts have been made to exploit other findings or claims with reference to the acceptability of advertising in various media, but only a small part of this can withstand critical inspection.

The Stanton and Lazarsfeld Program-Analyzer and the Schwerin Reacto-Meter, currently being applied to the measurement of interest in radio programs, give some indication of the influence of the medium, since the broadcast program is the medium with which the commercial message is associated. Fundamentally, these two similar techniques enable large groups of people to indicate in detail the points which they like or dislike in a program. This opinion method of measurement can be broken down to broadcast periods of a few seconds duration. In fact, any point can be selected for study. The results are applied to improving a program to the maximum within the capabilities of the creative talents. They are not intended to provide a basis for comparing advertising attitudes toward different media.

SAMPLING FOR MEDIA RESEARCH

Most media research is based upon responses from samples of people intended to represent known population groups. Therefore, it is of the utmost importance that the samples be adequate for the particular problem to which it is applied. The *Life* studies, which are projected to total national population with exceptions noted in the reports, must be appraised partly on the basis of the adequacy of the population samples. Radio-rating services, if they are to become projectible, must meet the same criterion. In fact, all audience studies, whether national in scope, local, or selective as in the case of circulation lists, must meet with the requirements of proper sampling if they are to be considered valid.

A SAMPLE EVALUATION STUDY

Alfred Politz Research (1946) made an evaluation study of the sample used in McCall's *Qualitative Study of Magazines*. This demonstration compared the samples as selected in the three McCall studies with a sampling procedure of known precision developed by Alfred Politz. The essential difference in sampling methods lies in the fact that the regular samples for the McCall studies were based upon quotas assigned to the interviewers, but within which the interviewer was permitted to select the homes at which calls were made. The precision sample, on the other hand, allowed the interviewers no freedom of choice whatsoever in the selection of homes to be called upon. The precision sample had further characteristics, confirmed in the experiment, which indicate that it is free of the kinds of sampling bias which interviewers are suspected of introducing when they work on quotas.

The evaluation study revealed two very significant facts, relative to the sampling. First, it showed that the particular procedure used by McCall's gave results which implied that 40 per cent more of the homes owned a current issue of a leading women's service magazine than were known to exist within the population. The precision sample, on the other hand, produced results with no significant variance from the known circulation facts. Second, despite this startling revelation as related to audience measures, the attitudes elicited from readers within the precision sample were almost identical with those of the readers in the regular McCall sample. Since McCall's was making a study of attitudes the results gave confirmation to the three studies which had been conducted on the quota sampling method.

SOLICITED RESPONSES

SALES AND INQUIRIES

Most advertisements solicit an overt response, such as making an inquiry or purchasing the product. Self-addressed coupons, coded as to source, appear on many advertisements, though probably on fewer than 10 per cent of the full-page advertisements in certain general magazines. Advertising experience has shown that coupon returns are affected by the nature of the medium and by the audience reached. Starch (1927) found a significant relationship between coupon inquiries and sales. Obviously, any findings about the significance of coupons can be applied safely only to the small minority of advertisements which carry them. Certain characteristics of the advertising media in which they appear are also inferred.

THE PRODUCT-USER TECHNIQUE

Both *Life Magazine* (1940) and the Columbia Broadcasting System (1941) have published surveys correlating the ownership of advertised products with exposure to advertising messages in these media. The inference may be that if persons who have been exposed more to the messages of a particular advertiser also buy more of his products, then the purchase of the product can be credited to the particular advertisement or advertising medium. These published studies showed a very significant difference in possession of the advertised products as related to advertisement reading and program listening. The Columbia Broadcasting System claimed 53 per cent more ownership of the advertised products among listeners than among nonlisteners to their forty sponsored evening programs.

The product-user technique of advertising and media evaluation has been criticized because of the factors influencing sales which are not accounted for. Stanton, 1941, refined the product-user technique as applied to radio by equalizing the listener and nonlistener samples in many respects including exposure to advertising in magazines. The results showed less startling effects of the particular advertising in question, but Stanton's findings, like others on various media, showed strong productivity on the part of advertising. However, it is doubtful that all of the factors influencing sales have been accounted for in any study, and there are other considerations. For example, in dealing with printed advertising, it seems likely that ownership of a product may cause reading of advertisements just as logically as the reading of the advertisements may lead to the ownership of the product. This technique seems in need of further refinement.

VALUE OF ADVERTISING MEDIA

Advertisers and students can compare the values of different advertising media only when the units of measurement are comparable. One comparable measure is the size of the audience which can be psychologically impressed with advertising messages per dollar invested. This is not expected to bring all media up to the same fixed standard, but such information on all media would be a great step forward in the matter of making comparisons. In a more general way, the total audiences of the media themselves, when they are actually known for all major media, will help in making evaluations. It will also be valuable to know the composition of these audiences in terms of the various breakdowns having marketing significance.

Audience attitudes, while perhaps a lesser function of the advertising medium, are by no means unimportant to the advertiser. The medium provides the frame of reference in which an advertisement is seen or heard, and it may contribute something to the attitude aroused by the advertising message. It is only the problem of measuring these attitudes, not the question of their significance, which has given this aspect of advertising media so little emphasis. The future challenge lies far more in the appraising of attitudes than in the enumeration of audience size.

READING REFERENCES

The Advertising Research Foundation, *Continuing Study of Farm Publications.* New York: The Foundation, Study No. 1, 1946, pp. 61–64.

———, *Continuing Study of Transportation Advertising.* New York: The Foundation, Study No. 7, 1946, pp. 17–22.

Agnew, H. E., and W. B. Dygert, *Advertising Media.* New York: McGraw-Hill Book Company, Inc., 1938, Chap. 5, "Factors to Consider in Comparing Rates."

Duffy, Ben, *Advertising Media and Markets.* New York: Prentice-Hall, Inc., 1939, Chap. 3, "Choosing the Media."

Lazarsfeld, P. F., *Radio and the Printed Page.* New York: Duell, Sloan & Pearce, 1940, Part V, "Radio and the Printed Page as Sources of News."

———, and F. N. Stanton, *Radio Research,* 1942–43. New York: Duell, Sloan and Pearce, Inc., 1944, Part III, "The Program Analyzer."

Life Magazine, *Continuing Study of Magazine Audiences.* New York: Report No. 1, 1938, pp. 7–10.

———, *Continuing Study of Magazine Audiences.* New York: Report No. 8, 1946, pp. 9–12.

McCall Corporation, *Qualitative Study of Magazines.* New York: The Corporation, 1946, Part II, "The Evaluation Study."

56. CUSTOMER RESEARCH

JOHN G. JENKINS, PH.D.

Deceased. Formerly: Professor of Psychology and Chairman, Department of Psychology, University of Maryland; Captain, USNR and Chief, Aviation Psychology Section, Bureau of Medicine and Surgery, Navy Department; Chairman and Director of Research of Committee on Selection and Training of Aircraft Pilots, National Research Council.

During the first two decades of the twentieth century, text books of psychology often contained lists of instincts which were described as the moving forces behind all human behavior. In like fashion, the early books on psychology in advertising tended to feature lists of general appeals which were presumed to underlie all buying behavior. Even today, an occasional writer will imply that, to sell merchandise or services, one has only to appeal to sex, or hunger, or pride, or desire for economy.

General statements of this sort have never had much practical value for those who produce advertising. The motives listed possess undoubted strength but there is nothing to indicate which one will prevail in a given buying situation. The same buyer who will walk a mile to save three cents on canned goods may insist on F-Center seats in the theater. A person's desire for safety in an automobile may be entirely overlaid by his desire for speed and swank in his next car. It is necessary for the advertising agency to discover, as accurately as possible, which specific determinants govern the purchase of the products turned out by their clients. Merely to know that the public is susceptible to many different general motives is of little significance.

ANALYSIS OF BUYING BEHAVIOR

The key to future buying behavior, for any specific article, is to be found in the study of existing buying behavior with regard to that article. That statement is the basic thesis of customer research, a technology which has grown very rapidly from a small beginning in the early twenties. Today there are numerous organizations devoted entirely to customer research, and there are many advertising agencies and large retail stores which maintain their own customer research departments. Many, if not most, of these began with the assignment of a single individual, followed by a rapid expansion to a department of considerable size and importance. Today those who are in charge of distribution and publicity need not rely on general lists of human motives, for they can study at first hand the buying behavior connected with the sale of a specific commodity.

TECHNIQUES OF STUDY

Three main ways of studying buying behavior have been developed in customer research: (1) controlled inquiry, in which questionnaires and interviews are used to ask respondents about their buying habits and motives; (2) field observation, where the investigators go to the normal purchasing outlets under instruction to make specified observations as buying goes on under normal conditions; and (3) deliberate experiment, in which special conditions are established whereby it is possible to vary at will the specific factors under investigation. These methods have been subjected to successive improvements and refinements; and all aim at obtaining specific and accurate information regarding the behavior of a particular sampling of buyers. Just as the insurance companies can regulate premiums on the principle that accurate records of the past and present provide the key to what will happen in the future, the marketer has learned that he can establish sound marketing policies by analyzing the buying behavior of customers.

INQUIRY

Of the three methods, the use of the questionnaire and the interview suffers most from amateurish handling. A poorly planned questionnaire, lacking any proper pretest, and distributed to an inadequate sampling, can return amazingly inaccurate results. In the hands of experts who know its shortcomings and know how to compensate for them, on the other hand, the questionnaire has shown itself to be a powerful tool for obtaining information otherwise unavailable.

PREPARATION OF QUESTIONNAIRE

The formulation of an adequate questionnaire begins with a period of informal field study, which indicates what to ask, how to ask it, and generally sets up hypotheses for later investigation. Having thus obtained the 'feel' of the situation, the psychologist makes up a preliminary form of his questionnaire which he then subjects to a pretest. This most important stage calls for revisions where questions are found to be ambiguous, or vague, or loaded, or incomprehensible. It provides for changes in question-order where one question is found to influence the answers on a later question. Months may be spent in these preliminaries on a questionnaire which is finally administered in a few days.

Obtaining a Representative Sample. Unless the sample used in any survey accurately reflects the total group of potential buyers, the most carefully prepared questionnaire is very likely to produce tragically misleading results. Where the early research agencies turned unhesitatingly to the telephone book for lists of respondents, the designation of an adequate and reliable sample has now become a highly specialized task to be intrusted only to experts. The largest agencies frequently call in outside specialists to consult on sampling problems. The novice in customer research will do well to be guided by this practice, since the penalty for inadequate sampling is often the establishment of marketing policies which later events will prove to be unsound.

CONTACT WITH RESPONDENTS

Responses to questionnaires may be obtained by mail or by face-to-face interview. Mailing the questionnaire may appear economical, but the returns received are likely to represent such a small percentage of the addressees as to make this method more expensive than that of interviewing. In most cases interviewing is more desirable because it permits better control of the sample, obtains a much higher rate of returns, and provides for a wider range of information.

Training of the interviewers is vitally important, and close supervision of the interviewing process is essential. Above all, interviewers must be trained to report accurately, so far as possible recording the actual words of the respondent. Only in this way can incidental remarks be secured for use as guideposts in interpreting the numerical data. A spontaneous remark to the effect that the bottle for a certain product has an inconvenient shape may prove to be more important than all of the quantitative results obtained for the product contained in the bottle.

USING RESULTS

The most difficult stage of the task confronting the director of a consumer research is that of interpretation of results. If interpretation is based only upon inspection of gross tables it is quite likely that the most significant implications of the study may be missed. Thus in a tabulation for soaps it may be found that the national consumption for a given brand shows 75 per cent of the population buying flake soap with only 25 per cent buying bar soap. To end the study here is to miss the point entirely. A breakdown by geographic groups may show the South buying 50 per cent bar and the North buying only 15 per cent bar. Fractionation into income groups may reveal that the lower income groups buy four times as much bar soap as the higher income groups. Separation of old from new users may reveal strong loyalty among the older users for the original form of the soap. Distinction between those who buy in large quantities and the single-unit purchasers may show differences not even hinted at in the gross table. And this may be only a beginning of the breakdowns employed to get at the real meat of the study.

The director of consumer research pays great attention to the spontaneous remarks made by respondents and to the actual phrases used in replying to set questions. A chance remark by an older respondent, faithfully reported by the interviewer, may suggest that a breakdown by age will indicate significant differences. Careful consideration of the exact answers made to a given question may show that all of the answers cannot properly be classed together but must be separated into several groups, with comparisons to be made only within these groups. Above all, the investigator keeps his eyes wide open for information which he did not ask for, but which is important enough to force its way into the records.

Consumer research requires continual contact with the current researches on the questionnaire. By reference to articles published in the technical journals, the research director keeps abreast of new developments in the questionnaire method. In published researches, he finds definite information concerning limitations of various forms of the question, data on reliability and validity of questions, and evidence regarding the hundred-and-one aspects of the questionnaire method which have been subjected to careful investigation.

FIELD INVESTIGATIONS

One type of field investigation involves the study of the flow of customers. For example, a small portable recording polygraph is carried by the investigator, who records on a tape by pressure on the buttons of the instrument the exact path of a customer through the store together with the time spent in each location. The instrument is concealed under a book or package of some sort and as soon as the customer leaves the store, a new customer is unobtrusively picked up and followed. A collection of a number of such records makes it possible to draw up a flow-chart, showing most frequent paths followed and supplying information as to typical lengths of stay in various departments. The use of this technique in a commercial exposition showed that certain exhibits were inspected on the average only for a small fraction of the time consumed in their operation and it demonstrated at once that certain doorways must be blocked if the average visitor was to witness even a majority of the exhibits. The technique has untold potentialities for use in department stores and other outlets of the sort where the average path of travel is unknown and unconsidered in arrangement of space.

FIELD CREWS

A second type of field investigation involves the establishment of field crews to go out and check on the actual distribution of a commodity in retail outlets. A soft-drink manufacturer checked in this fashion on the actual temperatures at which his beverage was served. When his field crews found surprising variations in serving-temperature, the results were taken to indicate the necessity of making adequate refrigerating apparatus available to the outlets. A hair tonic manufacturer used a field crew to detect the sale of substitutes and to report on improper applications of his product. A manufacturer of flashlight batteries has sent out crews to learn about the age and operating condition of batteries in average households, and has modified his product in the light of these findings. A distributor who had been sending expensive window displays to retail outlets finally decided to employ a field crew to find out what actually happened to his displays. The results indicated a change in distributing policy which meant that fewer displays were routed directly to the waste bin or the home nursery.

RECORDING CONVERSATION

Modern electric recording equipment makes it possible to install concealed microphones and to record precisely what the customer says in retail outlets as a guide to future sales strategy. Where this cannot be done, shorthand notes can be obtained as a record of exact comments. Analysis of these records will show what hundreds of customers spontaneously remark about a given product. If 45 per cent of the women visiting automobile salesrooms make some remark about the upholstery of a car, while only 1 per cent show any concern about its horsepower, there is a strong suggestion that upholstery should be featured in advertisements. The same technique makes it possible for a salesman to review his own sales talks and to see objectively his own shortcomings. Where the technique has been properly employed, it has met with enthusiasm on the part of those to whom the records were made available.

EXPERIMENT

Deliberate experimentation returns the most dependable results of all methods of studying buying behavior, but it has not yet seen widespread use. In contrast to field observation, which leaves conditions as it finds them, deliberate experimentation involves a preconceived variation of specified conditions. Side by side comparison of drinks from which a specific element has been omitted yields final evidence as to the value of that element, providing order of presentation and nature of sampling have been properly controlled. Invitation to choose between colored and uncolored fountain pens brought quick assurance that the public would accept the colored pens. The planned shifting of a small counter selling standard merchandise at reduced prices has been used in a department store to study the value of different locations. The final result of deliberate experimentation is in direct proportion to the control of conditions and the observation of standard experimental precautions.

COMMERCIAL VALUE

Instead of offering general lists of motives which offer small aid in deciding which motive will operate in a given situation, psychologists today attempt to advise the marketer by analyzing actual buying behavior. Detailed studies of what purchasers actually do and say makes it possible to determine what to manufacture, what to feature in advertising the manufactured article, and where and how and when to attempt to sell it. There is no better guarantee of the validity of this technique than to point to the wide scale on which it is employed in present-day merchandising.

READING REFERENCES

American Marketing Society, *The Technique of Marketing Research.* New York: McGraw-Hill Book Company, 1937, pp. 62–92; 190–247; 283–346.

Bingham, W. V., and B. V. Moore, *How to Interview* (3rd ed.). New York: Harper & Brothers, 1941, pp. 162–179.

Blankenship, A. B., *How to Conduct Consumer and Opinion Research.* New York: Harper & Brothers, 1946, pp. 10–136; 241–288.

Jenkins, J. G., "The Questionnaire as a Research Instrument." *Transactions of the New York Academy of Sciences,* 1940, Ser. II, **2,** No. 5.

Lazarsfeld, P. F., "The Art of Asking Why in Marketing Research." *National Marketing Review,* 1935, **1,** 1–13.

Politz, A., "Can an Advertiser Believe What Surveys Tell Him?" *Printers' Ink,* April 5, 1946, pp. 23–25.

57. RADIO AND TELEVISION AUDIENCE RESEARCH

ERNEST DICHTER, PH.D.

Psychological Consultant; Independent Specialist in Market Research Advertising. Formerly: Columbia Broadcasting System.

The application of modern psychology to radio and television is of relatively recent date.

There are four major fields in which psychology has been able to provide valuable help solving problems confronting the broadcaster attempting not only to reach, but to satisfy, his listeners: (1) the radio audience, who and how many people listen; (2) psychophysics of listening, FM or AM, loudness, speed of delivery, etc.; (3) effects of radio listening, the impressions and ideas left with the listener; and (4) aesthetic problems.

The applied psychologist employed by a broadcasting company usually is confronted by the department head or the management, the producer or the advertiser, with one or more of these problems in the form of concrete and practical questions. His problem may be anything from investigating the influence on Frank Sinatra's success of a youthful audience's swooning and swaying, screaming, and other fanatical exhibitions, to finding what distinguishes morning news programs in psychological effect from the evening news, to studying the best way for Kate Smith to appear on her program—as the center of it or to appear only at the beginning and end of it, or to decide whether commercials are most efficacious at the beginning, middle, or end of a program, to determine which jingles are best remembered, to examine the harm done by the annoying effect of Lucky's repeated "Lucky Strike's green has gone to war," to conclude whether the "soap-opera" fan is different from other radio listeners and if so, in what respects, or to determine the psychological effect of daytime serials on the average housewife, to discover the best radio approach to the atomic bomb problem for the desired effect, and so forth.

RADIO RESEARCH TECHNIQUES

Many questions put to the psychologist can be answered by utilizing only relatively simple research technique. Sending out two thousand or so questionnaires related to a specific program to a carefully selected sample is adequate.

Other problems require the administering of personality tests to listeners and to nonlisteners of, for example, soap operas. In this way it can be determined if one group has characteristics distinguishing it from the other and what they are.

Certain questions are best investigated by inviting special listener panels to participate in specifically planned experiments. Through an arrangement where frequency and volume of sound can be regulated with the twist of a knob, for instance, the best liked frequency and volume has been determined. With a similar laboratory arrangement audience preference for tonal range performances was studied. This test (Eisenberg and Chinn, 1945) showed that a majority of listeners did not prefer a wide range. The investigators were led to conclude that under present conditions of sound reproduction the wide range is actually less pleasant than the narrow range.

There are also various technical devices existing in the field by which "like" and "dislike" reactions can be registered on a continuous strip of paper as the program is heard. The resulting curve gives the analyst a clear picture of the highs and lows of a program.

Uncovering just what goes on in people's minds as they listen to a program or watch a telecast is attacked by "depth interviewing," a special type of indirect associative questioning. When the purpose is to determine the actual change of attitude brought about by a broadcast, discussions are conducted before and after the program.

While the techniques of radio research vary with the problem posed, it seems to this writer that applied psychology's most genuine contribution to radio and television is the development of a methodology which can describe in detail, if not measure in a quantitative fashion, just what a radio program or a television show does to people, just what effect it has on them.

Therefore our concentration in this section is on recent experiments and the theoretical structure developed from them, to illustrate the application of modern psychology to such vital problems as what goes on in the listener's mind during and after a broadcast or telecast.

WHAT HAPPENS IN THE MIND OF THE LISTENER?

Everytime a person listens something registers in his mind—a feeling of pity, of hatred, an arousing or satisfaction of curiosity, horizons expanded, new frustrations added, a mood created or changed, or aesthetic senses awakened.

It is this inventory of a program's psychological effects that is of great importance to the producer and writer. The writer conceives a program with an effect in mind. Often he cannot be sure the script achieves its purpose. When the psychologist can go to the producer and report that during the second dramatic scene most listeners ceased to visualize the action, that they became psychologically blind, he is offering useful and valuable information. The producer knows that visualization is necessary for orientation into the program and the feeling of excitement created by it, and is able to remedy the situation by introducing appropriate techniques and language. In other words, concentration on the psychological effects of a program can give the producer a fairly accurate understanding of what makes his show click or what prevents it from doing so.

"EFFECT" DOES NOT MEAN "LIKE" AND "DISLIKE"

Analysis of "effects" of radio programs has to be clearly distinguished from "like" and "dislike" reactions to a program. In various studies done with radio listeners, movie-goers, television viewers, and theatrical audiences, when the subject was instructed simply to describe his feelings during and after the production and not asked for critical judgments, two major types of responses resulted.

EMOTIONAL REACTIONS. One group responds with a description of such effects—excitement, laughter, new insight—and expresses it with remarks such as: "Well, it made me feel funny," "It forced me to think of my own childhood," or "It put me in a good mood."

CRITICAL REACTIONS. The other group compares the production with set standards: "She didn't play as well as she should have," "There was nothing new in this program, only old jokes," "I knew after a few minutes what would happen in the end," or "It reminded me of a motion picture I saw twenty years ago."

What is the significance of this distinction? Just this: The psychological structures of these response types are entirely different; they compare with apples and pears. They cannot and should not be combined in final results.

When a listener states: "My heart beat faster," he is reporting an immediately observable phenomenon. It represents true scientific information. When another listener comments: "I didn't like his overacting" or "I thought it was a silly story," he is making an intellectual reflection or interpretation. Such critical statements are reliable only when they are accepted as indications of the fields from which the standards were taken. In other words, it is interesting that a listener chooses to introduce moral standards into his appraisal of a program. It may not be interesting or usable information that on the basis of this moral standard he found the show wanting. Ethical standards often lead to critical remarks.

In an experiment conducted with 80 radio listeners these two types of responses occurred in a ratio of about 85 per cent descriptive remarks to 15 per cent critical ones. Therefore, if an interviewing procedure of a mechanical or personal nature tends to evoke a much larger percentage of comparisons with critical standards, bias is being introduced into the research procedure.

One of the most important tasks for the applied psychologist in the field of radio and television is to help the broadcaster, producer, and advertiser to distinguish between whether his show has just been liked or actually has produced the desired effect.

Two methods have been developed to offer this type of psychological help to the practical broadcaster. *Content analysis,* that is, a psychological profile of the show in terms of gratifications offered, is one aspect; *depth interviews* with listeners during and after the program, is the second approach.

Both content analysis and depth interviewing can be used to weave a meaningful structure. The individual effects produced by the show and the mechanisms leading to these effects or preventing them from full functioning become clear. Understanding the effects produced by as important means of communication and mass education as radio and television rather than accepting like and dislike reactions often opens new and revolutionary avenues of use for radio and television. For example, it is generally accepted that the average radio listener does not like educational programs. Judging from this "expert knowledge" one could have assumed that television reactions would be about the same. Tests which the author conducted for the Columbia Broadcasting System applying psychologically sound methodology and concentrating on the gratifications, the feeling produced by the telecasts, proved this assumption to be very misleading.

In order to get reactions as natural as one would get in the respondent's home these tests were conducted in an atmosphere of comfort and informality. Essentially for this reason, only a few people participated in any one test. Programs from all New York television stations were used. The selection of programs was determined by what station was broadcasting at the time of the test. A total of 350 subjects were interviewed.

The length of viewing was not constant for all tests. Usual procedure was to interrupt viewing midway in a 2-hour broadcast, or at a time when films replace the live program. This intermission was used to discuss what had just been broadcast. No questionnaires were used, and few direct questions were asked. The author suggested topics, directed the discussion, and made sure that everybody got a chance to voice his opinions and ideas. Generally, the respondents were delighted to comment or criticize at great length.

A stenographic record of these sessions was taken unobtrusively. Thus reports on specific programs were received and turned over to the producer. As data accumulated, special topics were found to reappear in successive sessions and longer reports for the use of all CBS television producers were prepared.

Most of the people who participated in the weekly Columbia Broadcasting System's tests were looking at their first television program. A few had seen television before, but usually their experience was confined to a brief glimpse at the World's Fair or a theater lobby.

During the first few minutes of each session, most comments showed that the people got a kick out of just seeing that pictures can be flashed through the air. Soon, however, interest shifts from the phenomenon of television to the program being televised. Once the novelty of electronic pictures in the home was accepted, people wanted to know how good these pictures are going to be. They wanted to know what television will give them that they can't get through the radio or movies. In short, they seek gratifications in television that they can get through no other medium.

To predict television's success or failure on the basis of these "added gratifications" is perhaps placing too much emphasis on this phase. But there is little doubt that television must evolve as a completely new and independent medium of entertainment.

EDUCATION—AN IMPORTANT GRATIFICATION OF TELEVISION

Education is one of the gratifications that members of the experimental audiences seemed to anticipate in television—and here is one contradiction to "expert knowledge." Because of its possibilities for vivid, lively presentations, they expect television to perform a greater educational role than either radio or the movies. While movies have the same visual advantage, people usually go out to the movies in search of entertainment.

Most of the respondents felt that television would be particularly useful in showing them how to do things—how to become beautiful, how to play the piano, how to make a dress, how to use cosmetics. What they expect, in other words, is education that they can apply immediately for their own benefit. Since televiewing is a home activity, many practical lessons can be tried out immediately and this is probably one reason why so much interest in education television was found. Advertisers may well profit from this interest by providing "how-to-do" commercials.

Further contradicting the "known" fact—that radio shows do not succeed unless they are relatively short—was the desire of some viewers to see full-length operatic and theatrical performances. In other words, they want major entertainment values in return for the time. In these demands—or hopes—viewers frequently stressed the full-length character of such programs, thinking of them as substitutes for actual visits to the theater or opera house.

Thus concentrating on effects rather than on superficial likes and dislikes, new gratifications emerge for radio and television.

RESEARCH HELPS IN MAKING PROGRAM CHANGES

Controlled psychological experiments are also applied to radio and television research in an effort to discover needed changes in program.

In two years of analyzing soap operas, the characters appearing in these stories, the kind of troubles they get into, and the effect on the listeners, were studied in detail. The main attraction of soap operas was found to be that listeners consider the stories true to life and use them as guides for their own living, as lessons in living. This was true despite the fact that the majority of such stories are far from being true to life. The audience has such a great need for the type of advice offered that they are uncritical of the stories. Recognizing and accepting the need for advice and guidance, the researchers still were concerned because many of the lessons taught in these programs actually are unhealthy and unsound.

Rather than having the "blond" in the eternal triangular plot of soap operas eliminated by scheming, an attempt was made to rewrite the scene in such a way that the married couple saw that no blond could have a permanently disturbing effect on their marriage unless the marriage had somewhere been lacking in the first place. The couple was made to sit down and candidly discuss their difficulties and to arrive at the real causes of the rift in their marriage. Then they could proceed soundly.

Unfortunately broadcasting companies do not always look favorably upon such attempts to prove that psychologically sound radio programs are just as attractive to listeners as programs of lesser value. But whenever such experimentation was carried on, the writer was able to see the X ray of just what his program does to the listener.

The psychologist would be unrealistic if he denied that the superpractical radio man often doubts that the scientist has concrete research methods with which to X ray a program. The psychologist must be prepared to show that he can get answers from listeners that are not just opinions but actually projective clues beyond the influence of subjective and critical judgments.

What questions can be asked to give full information regarding the psychological functions fulfilled by a program for the listener? First of all, no matter how skillfully it is done, no progress is made by asking a respondent why he liked or disliked a program. Just assuming the simple factor of difficulty of comprehension—if a person does not understand a play he will not like it. But it will be very difficult to get him to admit that fact even to himself. Instead he will seek a rational explanation in the play itself. He will say that the plot is confused or that he could not concentrate on it because there were too many characters, etc.

Another factor which must be taken into consideration is that most typical audiences do not possess critical aesthetic attitudes. They often mistake shows for reality. While they do know that the piece is being performed by actors, at the same time they experience it as real.

The general interviewing problem can be compared to the one encountered by the diagnostician who must diagnose a disease about which very little is known. If he were to rely on statements made by the patient, the diagnosis would certainly not be a reliable one. Like the diagnostician the audience researcher has to construct first by the development of work models and psychological analysis a meaningful hypothesis about the function of a program. The listener does not know what the researcher suspects.

In psychological research the actual questions to be asked must be formulated each time and be specifically applicable to the program or the types of programs being tested. Since there are large groups of programs belonging to the same psychological type, the actual amount of work to be done in developing these questions is not too tremendous. Over a certain period of time a set of stock questions which can be used again and again is built.

For instance, where "trueness to life" seems to be a relevant reaction category, applicable questions can be inserted from the stock. Identification can be determined by asking about the vividness of specific characters. Mystery stories can be tested by finding at what point the listener forgets his actual surroundings and is completely absorbed in the story and at what point the story lets him slip back into his living-room reality.

Getting a clear picture of the feelings which accompany listening to different types of radio programs is in general the first task. In a stirring war story the question would be to determine how often, if at all, listeners felt pity for the described victims or felt revengeful toward the enemy. In an educational program the problem is to seek at what point the listener had a feeling of "Aha, that's why," and to uncover any material given in the program which he tried to memorize or of which he makes special mental note. The subjects could be asked: How often did you want the actor to repeat a passage, read it more slowly, or more clearly? Which scene or situation do you remember best? Did your mind wander? Did you become impatient? If so, when? Did you say to yourself, Oh, move along, I know that already?

All these questions give clues to the type of gratifications and the respective mechanisms at work that are helping or hindering the message which the program has to deliver.

PSYCHOLOGY APPLIED TO PROGRAM TYPES

COMEDY. Psychology can be applied systematically to various types of programs. An analysis on comedy programs, for example, proceeded in the following fashion.

The researchers first asked themselves: What is the psychological function of a radio comedy program? On the basis of content analysis they distinguished two types of comedy shows in answer to this question.

The first kind of comedy program has a personality, characters and gags, which forces the listener to admire the star performer. Typical of this type of program is Fred Allen, Charlie McCarthy, and Bob Hope. They are shrewd, smart, happy-go-lucky, and easy for the public to admire and proudly identify themselves with them.

The second type of program invites the listener to feel sorry for the performer or to identify himself in such a way that his own weaknesses become more endurable and somewhat glamorized. Star personalities of these shows are the frustrated type like Vera Vague, or the awkward type, such as Red Skelton and Eddie Bracken, or the dumb individual, Finnegan, Mortimer Snerd, etc.

Laughter is, of course, the main effect sought in a comedy program but an analysis of the two-fold process of identification shows that in order to produce laughter relief from tension must first take place. Helping the listener to gain distance to his problems by showing him his own misery befalling others, is the mainstay of the comedy program. The two-fold process of identification, that is, with superior or inferiors, is utilized by the producer to create insight in the listener into his own attitudes and thus to give him real help. Although he is admittedly only interested in making people laugh, the comedy script writer sees himself pushed into the role of a psychotherapist. The more his jokes help the listener, the more the listener laughs. Combining content analysis and psychological interviewing the psychologist is able to show the producer why gags failed to elicit laughter—the comedian was neither a clear target for superior nor inferior identification or the jokes did not deal with subjects applicable to the listeners' lives, and so forth.

MYSTERY. In studying mystery stories the radio psychologist proceeds similarly. The mystery story has three main functions: (1) intellectual function—brings order into chaos, everything works out (while in real life it doesn't always), clear thinking can reveal the problem, a vicarious achievement; (2) emotional function—first sharp increase of tension and then relief; and (3) far reaching function—everyday life leaves people emotionally starved but mystery stories offer excitement to bring about high tides in living.

Strange as it may seem, the desired end-effect of a mystery story is an increase of security—things work out and the listener can help solve the problem. But reaching an end-effect depends on the working of many mechanisms.

In mystery stories, for instance, there is the danger of too close an identification with the characters in the story. Freud shows that an individual who steps out in such moments declares himself out. In other words, if the story is too realistic the producer may lose his listener psychologically.

Personality problems also enter into this type of listening. Some people cannot keep enough distance; others are too aloof to become emotionally involved. In both cases the desired effect fails to take place.

But once the researcher knows that a feeling of security has to be the desired effect, he can spot where the producer lost his listeners' attention, when they were pushed back into their reality, and help him to avoid making the same mistake again.

DOCUMENTARIES. The psychologist plays an even more important role behind the documentary program. Often for this type of program he steps into the picture at an earlier stage. He can conduct a so-called "social diagnosis," telling the producer and the writer what he thinks should be the goal and effect produced by the show.

This type of problem is the task of a psychologist working on public service programs. For instance, he finds the average American living in the postwar years suffering from three types of frustrations: 1) political impotence plus a feeling that things in the country are in the wrong order, 2) a feeling of futility, and 3) lack of progress. These frustrations, as is the case with all frustration, lead to aggression. In a democracy the easiest target for aggression is the government. The citizen mentally visualizes an immense, intangible, unseen power sitting in Washington and guiding his fate. The result is that the citizen refuses to accept his responsibility; the "leader" is to blame for things, not himself. This situation is a challenge to the producer. The effects which must be wrought on the listener have been outlined. The producer asks the researcher how it can be done.

Modern psychological principles say that the answer is to stop escapism, and psychotherapy shows how to do it. The program must show the listener that his attitude is understandable, that at times all people feel the same way. It must take the listener by the hand, it must encourage him, it must not blame him.

Translation of these concepts into programs is the next step. Associating the message with an overpowering problem of actuality, like the Atomic bomb, is one way and did result in the CBS program "Operation Crossroad." The CBS series "Assignment Home" was another way of making these points. It put the message in the story and problems of veterans returning home. The translation has infinite possibilities and if the circumstances call for it, new types of programs are developed.

Test of the Effect. Once the program has been conceived and written, the final step is to test whether or not it actually has fulfilled the intended functions. Psychological radio research of audience reaction to a program involves the creation of insight, similar to the task of the psychotherapist. The psychologist's problem is somewhat easier than that of the therapist, however, because he does not necessarily have to create insight in the listener. It is sufficient if he understands what motivates the audience member.

The Psychologist Helps the Radio Producer. If the psychologist is able to help the program man pay more attention to what his program does to the listener, the producer and practical radio expert will get a clearer picture of the psychological functions his program is achieving. The report of the radio researcher will then become less of a critical whip to be accepted or rejected depending on whether it flatters or accuses the producer. Rather it will be a clinical instrument to provide the program department with an X ray picture and laboratory report of the psychological functions engendered in the listener.

This knowledge should enable the program man to do his creative work with more assurance and more sharply focused goals. And if it can be demonstrated to the writer, the producer, and the advertiser, that a show will be better accepted the more it produces certain effects—as greater security, better techniques of living, better adjustment to everyday life problems, psychological researchers will have contributed to setting high goals for radio and televison as a whole.

Thus psychology, applied in the important branches of mass communication, has truly become part of science "applied" to the pursuit of greater individual and democratic success.

READING REFERENCES

Cantril, H., *Gauging Public Opinion*. Princeton: Princeton University Press, 1944.

———, *Psychology of Social Movements*. New York: John Wiley & Sons, Inc., 1941.

Dryer, Sherman H., *Radio in Wartime*. New York: Greenberg Publisher, 1942.

Eisenberg, Philip, and Howard Chinn, "Tonal Range and Sound Intensity Preferences of Broadcast Listeners," *Proceedings of the Institute for Radio Engineers*, September, 1945, pp. 571–581.

Gallup, G. H., and S. F. Rae, *The Pulse of Democracy; The Public-Opinion Poll and How It Works*. New York: Simon and Schuster, Inc., 1940.

McIver, R. M., *Group Relations and Group Antagonisms*. New York: Harper & Brothers, 1944.

SUBJECT INDEX FOR VOLUME ONE

Accident proneness, 328
  occupational safety, 278
Accidents. *See* Accident proneness; Occupational safety; Street and highway traffic
Advertising. *See* Propaganda
Advertising media research, 364–371
  advertising audiences, 367
  audience attitudes, 369
  sampling for, 370
  size of audience, 365
  solicited responses, 370
  techniques of, 364
  value of media, 371
Alchemy, 43
Alcohol, 88
Analysis. *See* Item analysis; Job analysis
Animism, 42
Application blank, field salesman, 212
  scorable type, 212
  See also Biographical information
Aptitude and intelligence, 109–115
  guidance and self-guidance, 114
  measurements of abilities, 110
    aptitude for the professions, 111
    clerical aptitude, 111
    intelligence, 110
  tests in predicting vocational success, 114
Astrology, 42
Audiences, for advertising, 367
  attitudes of, 369
  size of, for advertising media research, 365
  See also Radio and television audience research
Autocoid substances, 89
Aviation, 341–348
  design of equipment, 345
  operations, 346
  selection, of aircrew other than pilot, 343
Aviation, selection, of pilots, 341
    commercial airlines, 343
    military services, 341
    private pilots, 343
  training and evaluation of proficiency, 344

Benzedrine, 89
Bio-mechanics, 311–322
  and industrial psychology, 319
  and industry, 316
    equipment and its operation, 316
    operator variability and automatic machines, 317
    use of psychological principles, 316
  legibility of dials, 313
  and military problems, 312
    communications, 314
    comparison of equipment, 312
    dial displays, 312
    environment, 315
    operation of controls, 313
    visual problems, 314
  training of human engineers, 321
Boredom and fatigue, 78
Building illumination, 283–288
  color of light, 285
  in factories, 287
  in homes, 286
  in libraries, 287
  nature of lighting, 284
  need for better, 284
  in offices, 287
  paint and good lighting, 286
  in schools, 287
  types, of illumination, 285
    of equipment, 284
  See also Illumination
Building ventilation, physiological factors in heat-balance equation, 291

Classification procedures of the military, Army, 239
  Navy, 244
Clerical aptitudes test, SRA, 201
Clerical personnel, 197–207
  job analysis, 198
  personality factors, 203
  selection by tests, 197
    SRA clerical aptitudes test, 201
    United-NOMA entrance tests, 201
    USES clerical test battery, 200
  tests, 198
    Minnesota clerical test, 200
    in practice, 202
    Psychological Corporation general clerical test, 201
  training of, 203
    business school practice, 203
    Dvorak-Dealey simplified keyboard, 205–207
    English usage, 204
    typists, 205
Cocaine, 88
Color, preferences, 353
Color blindness, degrees of, 309
  frequency, 310
  and job analysis, 310
  tests of, 310
Color vision, 309–311
  degrees of color deficiency, 309
  frequencies of color blindness, 310
  and job analysis, 310
  tests of color vision, 310
Communications, military problems of, 314
Comprehension, training in reading, 51
Copy testing, 355–363
  inquiry tests, 360
  readership (recognition), 355
  recall test, 357
  sales test, 362
  selection of tests, 362
Counseling, and employer-employee relations, 146
  of retail sales personnel, 216
Criteria, criterion scores, 164

Criteria, morale, 19
  and validity, 170–177
    criterion contamination, 174
    list of major criteria, 171
    weighting criteria, 176
Cross validation, 180
Cults, 45
Customer research, 371–374
  analysis of buying behavior, 372
  experimentation, 374
  field investigations, 373
  inquiry, 372
    questionnaire preparation, 372

Dictionary of Occupational Titles (DOT), 161
Driver tests, 327
Drugs and smoking, 87–90
  alcohol, 88
  autocoid substances, 89
  caffeine, 88
  opium, marihuana, and cocaine, 88
  strychnin and benzedrine, 89
  tobacco, 87
Dvorak-Dealey simplified keyboard, 205–207

Education, illumination in schools, 287
  and propaganda, 23
  and television, 377
  See also Radio and television audience research, television and education
Efficiency in work, 268–276
  incentives, 274
    complexity of, 274
    employee attitudes, 274
    other financial incentives, 276
    success of incentive wage system, 275
    wage payment systems, 274
  law of motion economy, 270
  micromotion techniques, 269
    chronocyclograph, 269
  motion and time study, 268
  standardizing and timing the job, 268
  training, 271
    principles of motor learning, 271
Emotional factors and fatigue, 82
Employer-employee relations, 146
  and industrial ventilation, 294
  and personal interviewing, 145
  See also Worker and manager, 263–267
Employment agencies, 256–260
  classification, 257
  job requirements, 256
  tested vs. nontested requirements, 257
  use of measures, 259
  worker analysis, 258
    interviews, 259
    tests, 258
Engineering, street and highway traffic, 333
Environment and fatigue, 82

Fabulous zoology, 42
Factor comparison job evaluation, 161
Fatigue and its alleviation, 74–84
  emotional factors, 83
  environmental factors, 82
  length of work period, 80
  measurement of, 75
    boredom, 78
    physiological aspects, 77
    work decrement, 75
  methods of work, 81
  nature of fatigue, 74
  personal factors, 82
  rest, 78
  sleep, 79
Fear, 128, 132
Federal civilian personnel administration, 229–236
  career system in government, 234
  Council of Personnel Administration, 233
  efficiency rating, 235
  Post Office study, 229
  research in engineering tests, 233
  stenographic and typing tests, 232
  tests of general adaptability, 232
Field salesman, 212–215
  application blank, scoring of, 212
Field salesman, interview of, 213
  testing, 214
Forced choice, and personality testing, 184
  and rating, 149
Forgetting, 124
Freudian influence in leadership, 4

General clerical test, Psychological Corporation, 201
Grading job evaluation, 160
Graphology, 45
Guidance. *See* Counseling
Guided learning, 120–126
  acquisition of skills, 122
  adaptive learning, 122
  direct observation, 123
  efficient learning, 122
  forgetting, 124
  guidance in, 122
  ideational learning, 124
  learning process, 120
  limitations to learning, 125
  manual guidance, 123
  motivation, 121
  outcomes of learning, 121
  transfer of training, 121
  variability of work, 123
  whole-part method, 123, 124
  work situation, 123

Hawthorne study (Roethlisberger), 146
Health and ventilation, 73
Human engineering. *See* Biomechanics
Hypnosis, 44

Illumination, 60–68
  acuity relations, 61
  brightness contrasts, 63
  brightness relations, 62
  chromaticity relations, 62
  contrast discrimination, 62
  distribution, 64
  factory work, 63
  glare, 64
  levels of, 61
  measures of light, 60
  optimum conditions, 68

Illumination, pleasant lighting, 66
    achromatic factors, 66
    chromatic factors, 67
  spectral quality, 65
    acuity, relation to, 65
    chromatic discrimination and, 65
    eye protection, 66
    injurious illuminants, 65
  time relations, 62
  variables of vision, 61
  visual defects, 63
  See also Building illumination
Incentives, and efficiency in work, 274
  reward-punishment (praise-blame), 97
Individual differences, among trainees, 227
  clerical aptitude, 111
  clerical personnel and personality factors, 203
  color vision, 309–311
  intelligence, measurement of, 110
  job incentives, 274
  measurement of abilities, 110
  motion and time study, 268
  operator variability and automatic machines, 317
  professions, aptitudes for, 111
  selection of aircrew, 341–343
  subnormal workers, 252
  worker aptitudes, 157–159
  See also Aptitude and intelligence; Bio-mechanics; Measurement
Industrial morale, 297–303
  the individual factors, 300
  major obstacles, 302
  measurement, 279
  See also Railroads; Morale
Industrial ventilation, 288–297
  complaint situations, 294
  conditions of efficiency, 294
  human temperature regulation, 293
  physical factors, 289
  physiological factors, 290
  thermal conduction curve, 290
  thermal gradient, 290
Intelligence. *See* Aptitude and intelligence
Interest in work and play, 100–108
  interest and adjustment, 108
  measurement of interest, 103
    interest inventories, 103–107
    interest tests, 107
  nature of interest, 100
    definitions, 100
  permanence of interest, 102
  types of interest, 101
Interest inventories, 103–107
Interviewing, of field salesman, 213
  morale, 21
  worker analysis, 259
  See also Personnel interviewing
Item analysis, 177–182
  alternative analysis, 181
  cross validation, 180
  difficulty analysis, 178
  item-criterion analysis, 180
  item-test analysis, 179
  rights analysis, 181
  test-criterion item analysis, 180

Job analysis, 135–142
  analyzing the job, 141
  application of, 135
    to increasing output, 137
    to job evaluation, 137
    to merit rating, 137
    to safety, 137
    to test construction, 135
    to training programs, 137
    to transfer and promotion, 137
    to vocational counseling, 136
    to vocational selection, 136
  definition of a job, 135
  and job analyst, 140
  performance of, 138
    definition of terms used in, 140
    forms for, 138–140
  technique for, 141
  See also Clerical personnel, job analysis
Job classification and evaluation, 160–162
  Dictionary of Occupational Titles, 161
  job evaluation, 160
Job classification and evaluation, methods, 160
    factor comparison, 161
    grading, 160
    point system, 160
    ranking, 160
  occupational classification, 161
  statistical studies, 161
  US Employment Service classification, 161

Kata-thermometer, 69

Law, of egocentrism, 265
    of induction, 266
Leadership and social effectiveness, 3–10
  achievement of social effectiveness, 7
  behavior vs. feelings, 5
  definition through tests, 3
  democracy and social effectiveness, 9
  economic social effectiveness, 7
  elements of social effectiveness, 6
  the Freudian influence, 4
  items of social effectiveness, 5
  leadership, 7
  marriage and social effectiveness, 6
  organized groups, 7
  personality vs. intelligence, 6
  the personality quotient, 6
  personality as social effectiveness, 6
  private practice, 8
  social ascendance and submission, 4
  social contact, 7
  social maturity, 5
  See also Morale, and leadership
Learning. *See* Guided learning
Legibility and typography, 56
Length of work periods, 80
Lighting. *See* Building illumination; Illumination
Likert sigma scales, 12

Management's use of personnel research in training, 220–228
  Army experience, 224

- Management's use of personnel research in training, individual differences among trainees, 227
  - realistic basis of training, 223
  - related to other personnel areas, 220
  - trainee placement and evaluation, 224
- Managers. *See* Workers and manager
- Marihauna, 88
- Marriage, and social effectiveness, 6
- Measurement, of boredom, 78
  - of effects of propaganda, 29
  - of fatigue, 75
  - of industrial morale, 297
  - of interests, 107
  - of morale, 19
  - scorable application blank, 212
  - of social attitudes, 13
  - of vocational and educational satisfaction, 130
  - of work decrement, 75
  - of worker interests, 155
  - *See also* Aptitude and intelligence
- Mechanical personnel, 208–212
  - nature of mechanical ability, 208
  - selection, 209
    - for routine occupations, 209
    - for skilled occupations, 210
  - training, 211
    - training within industry, 211
- Media research. *See* Advertising media research
- Memory, 124
- Metropolitan police personnel, 247–251
  - criteria for validation, 250
  - selection of, 248
  - testing, 248
  - training, 248, 249
- Micromotion techniques, 269
- Military morale procedures, 303–308
  - administration and, 306
  - morale, and the home front, 308
    - in US Army, 307
  - suggestions from other armies, 308
  - World War I, 304
- Military morale procedures, World War II, 305
- Military personnel management, 236–247
  - Army personnel procedures, 238
    - classification, 239
    - classification tests, 239
    - officer candidates, 243
    - officer selection, 243
    - selection, 238
  - beginnings, 237
  - Navy personnel procedures, 244
    - classification, 244
    - selection, 244
- Minnesota Clerical Test, 200
- Monotony. *See* Boredom and fatigue
- Morale, 16–23
  - conviction measurement, 19
  - criterion of, 19
  - determiners of morale, 17
  - industrial morale surveys, factors in, 18
  - interviewing, 21
  - and leadership, 22
  - measurement of, 19, 297
    - early inventories of, 20
    - and training, factors in, 17
  - participation measurements, 20
  - related areas, 21
    - community surveys, 21
    - public opinion polls, 21
  - and social security, 22
  - *See also* Military morale procedures
- Motion and time study. *See* Efficiency in work
- Motivation and self-direction, 93–99
  - drive of internal energy, 94
  - frustration, 98
  - goals of set, 95
  - incentives, 95
    - competition, 96
    - knowledge of results, 96
    - noise, 96
    - reward-punishment (praise-blame), 97
    - rhythm, 96
    - social incentives, 96
  - intention, 98
  - mental set, 93
- Motivation and self-direction, motor set, 94
  - persistence of set, 95
  - preparatory set, 95
  - prepared units of response, 94
  - self-direction in social contact, 99
- Motor learning. *See* Efficiency in work

- Numerology, 45

- Occupational safety, 277–283
  - accident-prevention techniques, 283
  - accident proneness, 278
  - age and accident rate, 278
  - annual cost of accidents, 277
  - blood pressure and accidents, 281
  - emotional adjustment, 280
  - length of service and accident rate, 278
  - mental ability and accident rate, 279
  - philosophies of accident prevention, 277
  - physical condition and accident rates, 279
  - sensory motor coordination, 280
  - sex and accident rate, 279
  - vision and accident rate, 282
  - working conditions and accidents, 282

- Performance evaluation. *See* Rating
- Personality and physique, 116–120
  - constitutional types, 117
  - indirect relations, 118
  - physiognomy and phrenology, 116
  - physique and crime, 118
  - psychosomatic medicine, 119
  - and racial type, 118
- Personality quotient, 6
- Personality tests and prediction, 182–191
  - biographical information, 187
    - reliability and validity of, 189
    - specificity of prediction, 190
  - forced-choice technique, 184
    - advantages of, 186

Personality tests and prediction, forced-choice technique, problems in, 185
  general problems, 188
  of occupational success, 184
  reasons for lack of success, 183
  Wherry's technique, 184
Personnel. *See* Mechanical personnel
Personnel administration. *See* Federal civilian personnel administration; Military personnel management
Personnel interviewing, 142–147
  and employer-employee relations, 145
    counseling, 146
    follow-up interviews, 146
    Hawthorne study (Roethlisberger), 146
    worker attitudes, 146
  employment interviews, 144
    interview record, 145
    stereotypes in, 144
  halo effect in, 145
  improvement of, 143
  interviewers, training of, 144
  other uses, 147
  sources of unreliability, 142
  valid uses of, 142
    limits of validity of, 143
Personnel research, in training, 220–228
Physical types and personality, 116–120
Physique. *See* Personality and physique
Play, interest in, 100–108
Point system job evaluation, 160
Policeman, 247–251
Prejudice and propaganda, 28
Process conception of psychological illness, 84–87
  commonness, 84
  correction of psychological illness, 86
  illness in general, relation to, 85
  in terms of learning and forgetting, 85
Product testing, 351–355
  color preferences, 353
  field studies, 353
  laboratory studies, 352
Product testing, shape preferences, 353
  tactile perception, 352
Production, and job analysis, 137
Professions, aptitudes for, 111
Projective techniques, and measurement of social attitudes, 14
Promotion and transfer, and job analysis, 137
Propaganda, 23–30
  and counter-propaganda, 28
  measuring effects of, 29
  prejudice in, 28
  and related social forces, 23
    advertising, 24
    education, 23
    morale, 25
  response of the public to, 28
  techniques of, 25
    band wagon, 27
    card stacking, 27
    common devices, 26
    essential pattern, 25
    glittering generalities, 26
    name-calling, 26
    "plain folks," 27
    testimonials, 27
  See also Psychological warfare
Pseudology, 41–46
  alchemy, 43
  animism, 42
  astrology, 42
  automatic authorship, 44
  cults, 45
  ectoplasm, 44
  fabulous zoology, 42
  graphology, 45
  hypnosis, 44
  numerology, 45
  pseudoscientific discoveries, 45
  spiritualism, 43
  subconscious movements, 44
  telepathy, 44
Psychological warfare, 31–36
  culturally conditioned weapons of, 31
  cumulative weapons of, 32
  postwar propaganda, 36
  psychology and force, 33
  relation of psychological and physical weapons, 31
Psychological warfare, techniques, 34–35
    instruments, 34
    methods, 35
  types of, 33
Psychosomatic medicine, 119

Radio and television audience research, 375–380
  and program changes, 377
  techniques, radio, 375
  television and education, 377
Railroads, 336–341
  marketing of transportation, 340
  morale, 339
  safety, 336
    in America, 337
    in Europe, 337
  selection, 338
    in America, 338
    in Europe, 338
  training, 339
    of supervisors, 339
Ranking job evaluation, 160
Rating, 147–154
  errors in, 153
    halo effect, 153
    leniency, 153
    personal equation, 153
  forced-choice technique, 149
  graphic rating scale, 149
  man-to-man comparison scale, 148
  reliability and validity of, 154
  Richardson-Kuder scale, 149
  safeguards in, 153
  types of scales, 148–153
    samples of, 150–152
Reading efficiency, 49–54
  comprehension of what is read, 51
  the good reader, 49
  increase of, 53
  psychological processes, 50
  purposes of, 49
  reactions to and applying what is read, 52
  speed, 52
Remmers master scales, 13
Repeaters and accidents. *See* Accident proneness

Research. *See* Advertising media research
Rest, 78
Retail sales personnel, 215–220
  counseling of, 216
  and labor market, 216
  psychological research and influences, 217
  selection, 216
  training, 217
Richardson-Kuder scale, 149
Roethlisberger (Hawthorne study), 146
Rosander behavior scales, 13

Safety. *See* Occupational safety; Street and highway traffic
Salesmen. *See* Field salesmen; Retail sales personnel
Sampling, for advertising media research, 370
Scorable application blank, 212
Selection devices, 192–194
  conventional techniques, 192
  cost estimation, technique for, 193
  criterion of success, 193
Self-direction. *See* Motivation and self-direction
Seniority and accident rate, 278
Shape, preferences, 353
Sleep, 79
Smoking. *See* Drugs and smoking
Social attitudes, 10–16
  of contemporary societies and social groupings, 14
  development of attitudes in the individual, 11
  measurement and standardized measures, 12
    analogous measurements, 13
    behavior scales (Rosander), 13
    master scales (Remmers), 13
    method, 12
    projective techniques, 14
    psychophysical scales (Thurstone), 12
    sigma scales (Likert), 12
    types of scales, 12
  stability of attitudes, 11
  value of attitude measurement to society, 14
Social effectiveness. *See* Leadership and social effectiveness
Spiritualism, 43
Street and highway traffic, 325–335
  driver characteristics, 325
    accident data analysis, 325
    accident proneness, 328
    medical tests and accident rate, 328
    motor coordination and skill, 326
    results from driver tests, 327
    road tests and driving skill, 326
    theory of accident causation, 330
  driver and pedestrian education, 330
  human factor, 325
  regulation and enforcement, 332
  traffic engineering, 333
Strychnin, 89
Subnormal workers, 251–256
  attitudes of, 253
  environmental influences, 254
  individual differences, 252
  intelligence and job success, 252
  learning problems, 255
  mental efficiency, 253
  motor abilities, 253
  nonverbal abilities, 252
  occupational intelligence levels, 252
  social effectiveness, 255
  training, 254
  work adjustments, 255
Success. *See* Personality tests and prediction
Supervisors. *See* Railroads, training of supervisors

Tactile perception, 352
Telepathy, 44
Television. *See* Radio and television audience research
Testing. *See* Product testing
Tests, color vision, 310
  driving skill, 328
  and effects of ventilation, 72
  Federal civilian personnel administration, 232
Tests, of field salesmen, 214
  of police personnel, 248
  worker analysis, 258
  See also Aptitude and intelligence; Clerical personnel; Clerical tests; Copy testing; Criteria and validity; Item analysis; Personality tests and prediction; Selection devices; Vocational test construction
Thermometers, 69
Thurstone psychophysical scales, 12
Time and motion study, 268
Tobacco, 87
Traffic. *See* Street and highway traffic
Training within industry (TWI), 211
Transfer. *See* Promotion and transfer
Typewriter keyboard (simplified), 205–207
Typography, 55–60
  legibility, 56
  recommended kinds, 58
  visibility, 56

Unemployment and psychology, 37–41
  general unemployment, 38
    education of social attitudes, 39
    occupational adjustment, 40
    self-guidance, aids to, 40
    social control, 39
  individual unemployment, 37
    attitudes, 38
    psychological services, 37
United-NOMA business entrance tests, 201
US Army. *See* Military morale procedures; Military personnel management
USES clerical test battery, 200

Validity of criteria, success, 193
  See also Criteria and validity
Ventilation, 68–74
  atmosphere and the individual, 73

Ventilation, effect on test and work performance, 72
  and feelings of warmth and comfort, 71
  and health, 73
  measures of, 69
    dry-bulb and wet-bulb thermometers, 69
    effective temperature, 69
    globe thermometer, 69
    Kata-thermometer, 69
  and physiological functioning, 69
  See also Industrial ventilation; Building ventilation
Visibility and typography, 56
Vocational and educational satisfaction, 126–132
  causes of satisfaction and dissatisfaction, 126
    ability and interest, 129
    behavior, 126
    fear, 128
    morale, 129
    satisfaction, 128
      cycles of, 130
    security, 128
Vocational and educational satisfaction, social prestige, 128
  increase of, 131
    by overcoming fears, 132
    through sources of satisfaction, 131
  measurement of, 130
    clinical interviews, 130
    observation, 131
    questionnaires, 130
    rating scales, 130
    sociometry, 131
Vocational test construction, 163–170
  criteria of proficiency, 168
    weighting of elements of, 168
  criterion scores, 164
  experimental test, 164
    analysis of, 165
    selection, 164
    trying out, 165
    validation of, 165
  standardizing, 166
    norms, 166
  techniques of construction, 163
  varieties of tests, 163
Vocations, career system in government, 234
  clerical aptitude, 111
  clerical personnel, 197–207
  clerical tests, 198
  guidance and self-guidance, 114
  tests and prediction of success, 114
  See also Counseling

Warfare. *See* Psychological warfare
Wherry's technique, 184
Work, interest in, 100–108
Work decrement, 75
Worker and manager, 263–267
  law, of egocentrism, 265
    of induction, 266
  motivation, 264
  practical problems, 263
  psychological problem, 264
Worker analysis, 155–159
  of aptitude, 157–159
  of interest, 155
  of physical capacity, 155
  of skills and knowlege, 156
    trade tests, 156
Workers. *See* Subnormal workers

Form No. 3.

PSY, RES.L-1

Bureau of Educational & Psychological Research Library.

The book is to be returned within the date stamped last.

20.9.60		
20 MAR 1961		
20 APR 1961		

WBGP-59/60-5119C-5M

OXFORD BOOK
& S.Co.
CALCUTTA
NEW DELHI

150.13
FRY
Form No. 4

BOOK CARD

Coll. No. 150.13
Author Fryer Douglas H.
Accn. No. 197
Title Applied psychology

Date.	Issued to	Returned on
20.9.60	K. C.	
15.1.61		Retd.
20 MAR	86	

150.13
FRY
V. 1